**1**  **Operational Hydrological Forecasting during the IPHEx-IOP Campaign –**

**2**  **Meet the Challenge**

3  Jing Tao[1], Di Wu[2,3], Jonathan Gourley[4], Sara Q. Zhang[2,5], Wade Crow[6], Christa

4  Peters-Lidard[7], Ana P. Barros[1*]

5  1. Dept. of Civil and Environmental Engineering, Duke University, Durham, NC

6  2. NASA/GSFC Mesoscale Atmospheric Processes Laboratory, Greenbelt, MD

7  3. Science Systems and Applications, Inc., Lanham, MD

8  4. NOAA/National Severe Storms Laboratory, Norman, OK

9  5. Science Applications International Corporation, McLean, VA

10  6. USDA-ARS, Hydrology and Remote Sensing Laboratory, Beltsville, MD

11  7. NASA/GSFC Hydrological Sciences Laboratory, Greenbelt, MD

12  **\*Corresponding Author:**

13  Dr. Ana P. Barros

14  E-mail: barros@duke.edu

15  Phone: +1 919 660 5539

16 **Abstract**

17     An operational streamflow forecasting testbed was implemented during the Intense

18 Observing Period (IOP) of the Integrated Precipitation and Hydrology Experiment (IPHEx-IOP)

19 in May-June 2014 to characterize flood predictability in complex terrain.   Specifically,

20 hydrological forecasts were issued daily for 12 headwater catchments in the Southern

21 Appalachians using the Duke Coupled surface-groundwater Hydrology Model (DCHM) forced

22 by hourly atmospheric fields and QPFs (Quantitative Precipitation Forecasts) produced by the

23 NASA-Unified Weather Research and Forecasting (NU-WRF) model.   Previous day hindcasts

24 forced by radar-based QPEs (Quantitative Precipitation Estimates) were used to provide initial

25 conditions for present day forecasts.   This manuscript first describes the operational testbed

26 framework and workflow during the IPHEx-IOP including a synthesis of results.   Second,

27 various data assimilation approaches are explored a posteriori (post-IOP) to improve operational

28 (flash) flood forecasting.   Although all flood events during the IOP were predicted by the IPHEx

29 operational testbed with lead times of up to 6 hours, significant errors of over- and, or under-

30 prediction were identified that could be traced back to the QPFs and subgrid-scale variability of

31 radar QPEs. To improve operational flood prediction, three data-merging strategies were pursued

32 post-IOP: 1) the spatial patterns of QPFs were improved through assimilation of satellite-based

33 microwave radiances into NU-WRF; 2) QPEs were improved by merging raingauge observations

34 with ground-based radar observations using bias-correction methods to produce streamflow

35 hindcasts and associated uncertainty envelope capturing the streamflow observations, and 3)

36 river discharge observations were assimilated into the DCHM to improve streamflow forecasts

37 using the Ensemble Kalman Filter (EnKF), the fixed-lag Ensemble Kalman Smoother (EnKS),

38 and the Asynchronous EnKF (i.e. AEnKF) methods.   Both flood hindcasts and forecasts were

39    significantly improved by assimilating discharge observations into the DCHM. Specifically,

40    Nash-Sutcliff Efficiency (NSE) values as high as 0.98, 0.71 and 0.99 at 15-min time-scales were

41    attained for three headwater catchments in the inner mountain region demonstrating that the

42    assimilation of discharge observations at the basin's outlet can reduce the errors and

43    uncertainties in soil moisture at very small scales. Success in operational flood forecasting at

44    lead times of 6, 9, 12 and 15hrs was also achieved through discharge assimilation with NSEs of

45    0.87, 0.78, 0.72 and 0.51, respectively. Analysis of experiments using various data assimilation

46    system configurations indicates that the optimal assimilation time window depends both on basin

47    properties and storm-specific space-time-structure of rainfall, and therefore adaptive, context-

48    aware, configurations of the data assimilation system are recommended to address the challenges

49    of flood prediction in headwater basins.

50

## 1. Introduction

55

56        Floods are the most ubiquitous natural hazard, and flashfloods in particular remain a

57    leading cause of natural hazard deaths in the US (NRC, 2005).   Due to rapid flow responses ($\leq 6$

58    hours) at small spatial scales and large uncertainties associated with all hydrometeorological and

59    hydrological processes involved in the forecasting chain, flashflood prediction remains a grand

60    challenge in operational hydrology (Collier, 2007), including Quantitative Precipitation

61    Estimates (QPEs) (Ciach et al., 2007; Gourley and Vieux, 2005; Kirstetter et al., 2012; Tao and

62    Barros, 2013; Vasiloff et al., 2007; Zoccatelli et al., 2010), Quantitative Precipitation Forecasts

63    (QPFs) (Amengual et al., 2009; Cuo et al., 2011; Davolio et al., 2013; Dietrich et al., 2009; Jaun

64    and Ahrens, 2009; Mascaro et al., 2010; Rabuffetti et al., 2008; Rossa et al., 2011; Zappa et al.,

65    2010), highly non-linear model representations of hydrological process (Garambois et al., 2013;

66    Garcia-Pintado et al., 2009; Zappa et al., 2011), and probability-based decision rules (Coccia and

67    Todini, 2011; Dietrich et al., 2009; Hersbach, 2000) or threshold-based (either for rainfall or

68    discharge level) warning criteria (Demargne et al., 2009; Martina et al., 2008; Norbiato et al.,

69    2008; Rabuffetti and Barbero, 2005; Welles et al., 2007) as well. The predictability of

70    flashfloods is particularly  challenging in ungauged/poorly gauged and remote basins (Moore et

71    al., 2006; Norbiato et al., 2008; Reed et al., 2007; Tao and Barros, 2013; Versini et al., 2014) and

72    in mountainous regions where other geo-hazards such as landslides (e.g. debris flows) are often

73    associated with heavy rainfall (Band et al., 2012; Casadel et al., 2003; Liao et al., 2011; Tao and

74    Barros, 2014a; Wooten et al., 2008).

75        Operational hydrological forecasting and nowcasting for flashflood warning is stipulated

76    on  three tenets (Cloke and Pappenberger, 2009; Cuo et al., 2011; Droegemeier et al., 2000;

77    Hapuarachchi et al., 2011; Liu et al., 2012; Pagano et al., 2014; Vrugt et al., 2006): **1)**

78  availability of accurate QPFs with adequate lead times for effective warning and emergency

79  response; **2)** availability of  near real-time comprehensive observing systems (a variety of data

80  and observing systems, hereafter referred to as data support including ground- and satellite-based

81  QPEs, raingauge observations, and river discharge observations; and **3)** data assimilation systems

82  (DAS) to  merge and integrate available observations (i.e. discharge, satellite-based soil

83  moisture, etc.) into hydrologic models to improve initial conditions for flood forecasting using

84  physically-based distributed hydrologic models. Here, we briefly review each element and

85  propose strategies to improve the predictability of flashfloods in regions of complex terrain in the

86  context of the operational hydrological forecasting testbed implemented in the Southern

87  Appalachians for the Integrated Precipitation and Hydrology Experiment (IPHEx) campaign

88  (Barros et al., 2014). The use of physically-based and fully-distributed hydrologic models for

89  flood forecasting poses additional challenges on account of high nonlinearity of rainfall-runoff

90  response in space and time, further compounded by surface-groundwater interactions (Pagano et

91  al., 2014; Werner et al., 2009), which is also examined here with the Duke Coupled surface-

92  groundwater Hydrology Model (DCHM).

93      **1) QPFs** – Over recent years, ensemble prediction systems (EPS) for ensemble

94  streamflow prediction (ESP) have become increasingly ubiquituous in flood forecast operations

95  (Cloke and Pappenberger, 2009; Schaake et al., 2007), including the EFAS (European Flood

96  Alert System, Europe) (Alfieri et al., 2014; Bartholmes et al., 2009; Pappenberger et al., 2015;

97  Thielen et al., 2009), the operational HEPS (Hydrometeorological Ensemble Prediction System,

98  Switzerland) (Addor et al., 2011), and many others (De Jongh et al., 2012; Hsiao et al., 2013;

99  Nester et al., 2012; Pappenberger et al., 2015; Taramasso et al., 2005; Verbunt et al., 2007;

100 Zappa et al., 2010).  In the United States, the NWS's Hydrologic Ensemble Forecast Service

101  (HEFS), a part of the Advanced Hydrologic Prediction Service (AHPS) (Connelly et al., 1999;

102  Hogue et al., 2000; McEnery et al., 2005), operationally provides ensemble flow forecasts using

103  ensemble mean QPFs from multiple NWP models for flood risk management and other water-

104  related needs (Demargne et al., 2014).   However, NWP-based QPFs have long been found

105  inadequate in terms of rainfall intensity and variability, with cumulative rainfall amounts that

106  dominate forecast errors and uncertainty, especially for small to medium size basins and in

107  mountainous regions (Amengual et al., 2008; Cuo et al., 2011; Ebert, 2001; Jasper et al., 2002;

108  Lu et al., 2010; Pappenberger et al., 2005; Xuan et al., 2009). In addition, a gap exists among

109  meteorological operational practices for QPF and hydrological needs in terms of inconsistent

110  spatial and temporal resolution, approaches to bias correction and model output statistics (MOS),

111  and distinct points of view regarding validation and uncertainty (Demeritt et al., 2013;

112  Pappenberger et al., 2008; Shrestha et al., 2013). One advantage of the IPHEx operational

113  hydrological forecasting testbed is the seamless transfer of NWP QPF to the hydrological model

114  due to careful a priori planning and integration of the NU-WRF (NASA-Unified Weather

115  Research and Forecasting) and DCHM model requirements.

116      **2) Data Support** - Many campaigns, projects, and community workshops have been

117  devoted to improving the state-of-the-science and the state-of-the-practice of flood forecasting

118  (Amengual et al., 2008; Benoit et al., 2003; Davolio et al., 2009; Rotach et al., 2012; Schaake et

119  al., 2007; Zappa et al., 2008). Often, however, access to observing systems and data delivery

120  infrastructure, that is the data support, is lacking or remiss in terms of spatial and temporal

121  sampling density and extent, data quality and latency (Pagano et al., (2014).  The IPHEx testbed

122  was implemented in an environment with unique data support: 1) an extended observation period

123  (EOP) from October 2013 through October 2014 including the deployment of a science-grade

124 raingauge network of 60 stations (in place since 2007), half of which are equipped with multiple

125 raingauge platforms (during the IPHEx EOP, 2013-2014), in addition to the fixed regional

126 observing system including a disdrometer network consisting of twenty separate clusters, and

127 two mobile profiling facilities including MRRs (Micro Rain Radar); and 2) an Intense Observing

128 Period (IOP) from May-June of 2014 (IPHEx-IOP) focusing on 4D mapping of precipitation

129 structure during which NASA's NPOL S-band scanning dual-polarization radar, the dual-

130 frequency Ka-Ku, dual polarimetric, Doppler radar (D3R), four additional MRRs, and the

131 NOAA X-band dual polarized (NOXP) radar were deployed in addition to the long-term fixed

132 instrumentation (Barros et al. 2014). Like-minded field  campaigns, such as HyMeX

133 (Hydrological cycle in the Mediterranean Experiments)(Drobinski et al., 2014; Ducrocq et al.,

134 2014; Ferretti et al., 2014) and IFLOODS (Iowa Flood Studies) (Petersen and Krajewski, 2013),

135 focused on improving QPE for flood forecasting. The real-time ensemble hydrological

136 forecasting were conducted during the Special Observing Period of HyMex paying special

137 attention to uncertainties associated with QPF and its propagating along the hydrometeorological

138 chain and meanwhile advocating the consideration of uncertainties associated with initial soil

139 moisture and hydrological models as well[1](Vincendon et al., 2014) .  During the IPHEx-IOP, all

140 the data from deployed instruments, along with real-time discharge observations and the

141 operational radar-based QPE products (i.e. NSSL Q3 and NCEP/EMC Stage IV; see Section

142 2.2.2. for detailed description) were assembled together for operational hydrological forecasting

143 for the first time, and for synthesis and analysis a posteriori.

144    **3) Data Assimilation** – Even with the "perfect" hydrologic model and an "optimal"

145 combination of QPFs, QPEs and other data support, flood predictability depends heavily on the

---

[1] http://presentations.copernicus.org/EMS2014-461_presentation.pdf

146    realistic representation of  initial hydrological conditions (Berthet et al., 2009; Li et al., 2009;

147    Pagano et al., 2014). Data assimilation has proven an effective technique to reduce error and

148    uncertainty in initial conditions (as well as accounting for model errors) in flood forecasting

149    (Castaings et al., 2009; Komma et al., 2008; Madsen and Skotner, 2005; Noh et al., 2014;

150    Randrianasolo et al., 2014; Salamon and Feyen, 2009; Schaake et al., 2007; Vrugt et al., 2006;

151    Wanders et al., 2014;   among others), and in particular by assimilating available discharge

152    observations into hydrologic models (Bloschl et al., 2008; Clark et al., 2008; Lee et al., 2011; Li

153    et al., 2015; Li et al., 2014; Rakovec et al., 2012; Seo et al., 2003). However, the application of

154    data assimilation techniques to fully-distributed hydrologic models is still relatively rare due to

155    high nonlinearity and the large number of hydrological states (number of degrees of freedom)

156    involved (Lee et al., 2011; McLaughlin, 2002; Xie and Zhang, 2010), and the complex

157    implementation that requires correctly representing tempo-spatial uncertainty in forcing, model

158    parameters and structures, and observations as well (Clark et al., 2008; Crow and Reichle, 2008;

159    Crow and Van Loon, 2006; Flores et al., 2010; Noh et al., 2014; Ryu et al., 2009). Consequently,

160    a small number of studies are reported in the literature for real-world events (many are synthetic

161    studies), and even fewer for realistic operational flood forecasting (Liu et al., 2012; Rakovec et

162    al., 2015; Randrianasolo et al., 2014).  In this work, the impact of coupling the DCHM with a

163    river discharge DAS on the quality of both streamflow hindcasts and forecasts was examined in

164    the post-IOP phase of IPHEx. DAS experiments were conducted for different watersheds by

165    assimilating the discharge observations at the basin outlet using various techniques including the

166    EnKF (Ensemble Kalman Filter) (Evensen, 1994; Evensen, 2003), the fixed-lag EnKS

167    (Ensemble Kalman Smoother) (Evensen and van Leeuwen, 2000) and  asynchronous version of

168    EnKF (AEnKF) (Rakovec et al., 2015; Sakov et al., 2010). The testbed performance sensitivity

169 to the DAS configuration with regard to length of assimilation time windows (TW) and

170 assimilation frequency (AF) was also investigated for different basins.

171 This manuscript first describes the operational hydrological forecast activities during the

172 IPHEx-IOP in Section 2, and summarizes the real-time operational results during the campaign

173 in Section 3. Post-IOP analysis and synthesis, including the impact of implementation of data-

174 assimilation are presented in Section 4 with a focus on demonstrating the utility and added value

175 of the proposed strategies for improving flood forecasting in regions of complex terrain.

176

## 177 2. Operational Hydrological Forecast Implementation

### 178 2.1 Workflow of the Daily Operational Forecast

179 IPHEx was the first Ground Validation field campaign conducted in support of the

180 Global Precipitation Measurement (GPM) satellite mission after the launch of the core satellite

181 (Barros et al. 2014). The main objective was to characterize warm season orographic

182 precipitation regimes, the relationships among precipitation regimes and hydrologic processes,

183 and to investigate operational flashflood predictability in regions of complex terrain. The study

184 region is centered in the Southern Appalachians and spans the Piedmont and Coastal Plain

185 regions of North Carolina (Figure 1), with a focus on 12 headwater basins in the Southern

186 Appalachian Mountains (SAM) with drainage areas ranging from $71km^2$ to $520 \ km^2$ (Table 1).

187 The operational hydrological forecasting testbed during the IPHEx-IOP was conducted

188 collaboratively by Duke University (Duke) and NASA GSFC (Goddard Space Flight Center) to

189 issue 24-hour forecasts daily starting at 12:00 UTC for each one of the 12 headwater basins. In

190 practice, latency in the operational environment was constrained by computational resources and

191    the rates of data transfer from weather prediction at GSFC to hydrological prediction at Duke,

192    and thus the actual forecast lead time did not exceed six hours during the IOP.

193            Figure 2 depicts the operational workflow at Duke University to produce the daily

194    hydrological forecasts and hindcasts during the IPHEx-IOP (Barros et al. 2014). Specifically, 24-

195    hr forecasts provided by the NU-WRF model at GSFC were delivered to Duke daily around

196    8AM EDT.  The forecast fields were then projected into the IPHEx grid system (UTM17N) at

197    1km spatial resolution, interpolated to 5-min time-steps, and then converted into the format

198    required by the input interface of DCHM.  Multiple QPEs including Stage IV and Q3 for the

199    previous day were downloaded and processed on a daily basis to produce streamflow hindcasts

200    and provide updated initial conditions for the present day forecast.  The hindcast results were

201    evaluated for the 12 forecast points using previous day discharge observations downloaded daily

202    from the USGS (United States Geological Survey) online data portal.  In addition, the discharge

203    observations at the end of the previous day were nudged into the DCHM as the initial discharge

204    for the current day forecast, and the initial flow rates in channel pixels within each basin were

205    adjusted proportionally to the ratio of estimated streamflow to the observation at basin outlet.

206    The operational modeling system was implemented using MPICH2 (Message Passing Interface)

207    so that the operational forecast results, including streamflow forecasts for the present day and the

208    streamflow hindcasts for the previous day, could be produced every day before 3PM EDT.  Note

209    the operational system here was designed as such to mimic the timeline and overall framework

210    of the operational forecasting system at the National Weather Service River Forecast Centers

211    (RFCs), but actual public forecasts were not issued although it could be and results were posted

212    online at iphex.pratt.duke.edu. The ultimate goal of this study is to enhance the hydrological

213 forecasting skills through various strategies with minimum manual supervision and rescue as

214 needed in realistic operational systems.

215

216 **2.2 Hydrometeorological Forcing Fields**

217 **2.2.1 Quantitative Precipitation Forecasts (QPFs) and other atmospheric forecasts**

218 During the IPHEx-IOP, the NU-WRF operationally provided high-resolution 2D

219 forecasts of atmospheric forcing to drive the DCHM, including QPFs, air temperature at 2m, air

220 pressure at 2m, specific humidity at 2m, and wind speed at 10m, incoming shortwave radiation

221 and incoming longwave radiation at surface. The NU-WRF was implemented with 60 vertical

222 layers and three horizontal domains at resolutions at 9km (domain 1), 3km (domain 2), 1km

223 (domain 3) and 30sec temporal resolution. The model precipitation and atmospheric forcing

224 fields were output at 1km resolution and 5min intervals. Figure 3 shows the three horizontal

225 nested grids implemented in NU-WRF and the IPHEx domain. The NU-WRF physics

226 configuration include the Goddard 4-ice Microphysics scheme, the Grell-Devenyi ensemble

227 cumulus scheme, the Goddard Radiation schemes, the MYJ (Mellor–Yamada–Janjic) planetary

228 boundary layer scheme, the Noah surface scheme and the Eta surface layer scheme. The output

229 from the GFS (Global Forecast System) model every six hours at $0.5^{o}$ resolution were used as

230 initial and boundary conditions for the NU-WRF forecasts. More information about the NU-

231 WRF can be found in (Matsui et al., 2014; Peters-Lidard et al., 2015; Shi et al., 2014; Zaitchik et

232 al., 2013).

233

### 2.2.2 Quantitative Precipitation Estimates (QPEs)

During the campaign, two conventional ground-radar QPEs were used for operational hindcasts, namely Stage IV and Q3 data. An experimental ground-radar based QPE derived from the NOAA NSSL (National Severe Storms Laboratory) X-band dual-Polarized Mobile Radar (NOXP), and a satellite-based QPE, i.e. the NASA **I**ntegrated **M**ulti-satellit**E R**etrievals for **G**PM (IMERG), were also utilized for case studies after the IPHEx IOP. During the IOP, the operational QPEs (i.e. Stage IV and Q3) for the previous day were downloaded first, and then were (re-) projected to the IPHEx reference gridding system (i.e. UTM17 at WGS84). Q3 QPEs were resampled to the IPHEx common grid at 1km using the nearest neighboring method. Stage IV data were downscaled to 1km using a transient multi-fractal downscaling method (Nogueira and Barros, 2014). Details about each QPE are provided below.

a) Stage IV (Operational Radar-based QPE) - NCEP/EMC (Environmental Modeling Center) Stage IV data is a national multi-sensor 4km gridded hourly precipitation analysis with very short latency (about 1hour) (Lin and Mitchell, 2005). The Stage IV product is constantly updated with new analyses from the RFCs (River Forecast Centers), and the final product is available with a latency of 12~18 hours.

b) Q3 (Operational Radar-based QPE) - The Q3 or MRMS (Multi-Radar/Multi-Sensor) product provided by the National Mosaic and Multi-sensor QPE (NMQ) system at NSSL is a real-time nation-wide seamless QPE product at very high spatial (~1 km) and temporal (2 min) resolution which ingests rain gauge observations and hourly analyses of RAP (Rapid Refresh model) on the basis of 3D volume scan data from Weather Surveillance Radar-1988 Doppler (WSR-88D) network (Zhang et al., 2014). During the IPHEx-IOP, the hourly radar-based product with bias correction was operationally used for hindcasts. The 2-min radar-alone

12

257  products without gauge correction were also obtained after the campaign and used for analysis.

258  The Q3 is a real-time product, and thus its latency is on the order of minutes.

259      c) NOXP (Experimental Radar-based QPE) - The NOXP radar was deployed in the

260  Pigeon River Basin (shown in Figure 1) during the IPHEx-IOP (Barros et al. 2014). The radar

261  was installed at intermediate elevation (1176m) in the inner region, and operated with scanning

262  frequency of about 5 minutes and multiple sweeping elevation angles (from 0.5 to 8 degree),

263  which allows an unimpeded view for low-level across most of the inner basin to avoid terrain

264  blockage and overshooting, which are severe problems impeding the applications of

265  conventional weather radars in topographically complex terrain. Details about the NOXP radar

266  can be found in  Palmer et al. (2009).  Hybrid gridded NOXP data were produced by choosing

267  the lowest elevation angle without terrain blocking for each azimuth. The processed NOXP data

268  were gridded into UTM17 directly at the DCHM simulation resolution (i.e. 250m×250m) from

269  the radar-scanning spherical polar coordinate system. The algorithm components used in the

270  NOXP data processing (i.e. calibration, ground clutter removal, attenuation correction, DSD

271  retrieval, and QPEs, etc.) are described in (Anagnostou et al., 2013; Kalogiros et al., 2013a;

272  Kalogiros et al., 2013b; Kalogiros et al., 2014).

273      d) IMERG (Experimental Satellite-based QPE) - The IMERG Level 3 half-hour

274  precipitation products at $0.1^{\circ}$ x $0.1^{\circ}$ (Final Run) were used for the case studies in the post-IOP

275  phase of the campaign. The IMERG system integrates prior multi-satellite algorithms from

276  TMPA (TRMM Multi-Satellite Precipitation Analysis), CMORPH-KF (CPC Morphing –

277  Kalman Filter), and PERSIANN-CCS (Precipitation Estimation from Remotely Sensed

278  Information using Artificial Neural Networks – Cloud Classification System) (Huffman, 2015).

279  Specific details regarding the rainfall retrieval algorithm and the data (post)processing are

280    described in the Algorithm Theoretical Basis Document of IMERG (Huffman et al., 2014).

281    Similar to StageIV, the IMERG data were also downscaled to 1km using the fractal downscaling

282    method (Nogueira and Barros, 2014a and 2015).

283

### 284    2.2.3   Soil properties and historical hydrometeorological datasets

285    In preparation for the operational hydrological forecasting testbed, long-term historical

286    hydrometeorological datasets (atmospheric forcing and landscape attributes) necessary to

287    implement and operate hydrologic models in the Southeast US (shown in Figure 1a) at the

288    IPHEx reference resolution (hourly time-step, 1km×1km in UTM17N at WGS84) were

289    developed for a 7-year period (2007-2013), and are available on http://iphex.pratt.duke.edu. The

290    atmospheric forcing fields were downscaled from the North American Regional Reanalysis

291    (NARR) product with cloudiness-, elevation- and topographic correction (Tao and Barros,

292    2014c). The landscape attributes were constructed from MODIS land products by removing

293    cloud contamination (Tao and Barros, 2014b). Soil properties, including saturated hydraulic

294    conductivity, porosity, field capacity and wilting point, were extracted from the State Soil

295    Geographic (STATSGO) dataset[2]. Historical landscape attributes in the same day-of-year in a

296    wet year (2009) were used throughout the entire IPHEx-IOP period due to the lack of updated

297    MODIS products.

298

---

[2] http://iphex.pratt.duke.edu/DataCenter/Time-invariantDatasets/SoilParameters

14

**2.3 Duke Coupled surface-groundwater Hydrology Model (DCHM)**

299

300     The DCHM, implemented at 250m×250m spatial and 5min temporal resolution, was the

301     hydrologic model used for the operational hydrologic forecasting testbed. The DCHM is a

302     physically-based and fully-distributed hydrologic model solving water and energy balance

303     equations with coupled surface-subsurface interactions. Earlier studies using evolving versions

304     of the DCHM (formerly referred to as LSEBM, 1D-LSHM, and 3D-LSHM) were described in

305     various publications (Barros, 1995; Devonec and Barros, 2002; Garcia-Quijano and Barros,

306     2005; Gebremichael and Barros, 2006; Kang et al., 2013; 2012a; 2012b; Tao and Barros, 2014a;

307     2013; Yildiz and Barros, 2005; 2007; 2009) with demonstrated success particularly in flash-flood

308     and landslide prediction at event scale in the Pigeon River Basin (one of the core basins in this

309     study) (Tao and Barros, 2014a; Tao and Barros, 2013).  Before the IPHEx-IOP, the DCHM was

310     reinitialized and spun up (repeating simulations several times until internal equilibrium is

311     reached) for five weeks (April 1-May 5, 2014) driven by the ensemble of fractally downscaled

312     QPEs generated from the Stage IV product and historical hydrometeorological datasets in the

313     same month of a wet year (2009). Spin-up was conducted repeatedly until the flow difference

314     between the last and the current iteration is very small, i.e. the hydrologic system reaches internal

315     equilibrium, resulting in small stable simulated streamflow residuals. The final hydrologic states

316     at the end of the spin-up period were used as the initial conditions for the operational forecasts

317     starting on May 5. Note there was no tuning of initial conditions for the daily forecasts past May

318     5, and the model is uncalibrated.

319     The spatial and temporal resolutions of standard IPHEx products including NU-WRF

320     forecasts are respectively 1km and hourly.  All the forcing data were spatially interpolated to

321     250m using the nearest neighbour method, and landscape attributes data were linearly temporally

322    interpolated to 5min resolution. During the IOP, operational hourly QPEs (i.e. StageIV and Q3)

323    and 30min satellite-based QPE data (i.e. IMERG) were uniformly interpolated to 5min assuming

324    constant rainfall intensity, thus generally underestimating heavy rainfall intensities and

325    overestimating light rainfall (Nogueira and Barros, 2015) at times. NOXP QPEs (rainfall rate) at

326    radar scanning temporal resolution were averaged to 5min. Temporal interpolation of

327    atmospheric forcing fields including QPFs provided by NU-WRF was unnecessary since all the

328    fields were available at 5min resolution.

329

## 3.    Operational Results during the IPHEx-IOP

**3.1 Overview of the Operational Hydrologic Forecasting Testbed**

332    The overall forecast and hindcast results for selected headwater basins during the IPHEx-

333    IOP period (May 1 – June 15, 2014) are summarized in Figure 4. The QPFs provided by NU-

334    WRF overestimate rainfall for all twelve basins during the campaign, consequently

335    overestimating streamflow but capturing well peak times for all basins.  There were no missed

336    events, though several false alarms resulted from incorrect placement of rainfall cells in NU-

337    WRF QPFs (e.g. Basin 1 and 10).  The overestimation error is particularly large for the major

338    IOP event on May 15 in all basins, and for the secondary event on June 12/13 in the headwater

339    catchments of the Upper Catawba and Upper Yadkin (i.e. Basins 8-12, not shown here but can be

340    found on IPHEx website[3]). Some extraordinary flow forecasts (false alarms) are shown for May

341    30 in Basin 1, and on June 1 in Basins 4 and 5 which are attributed to the incorrect placement of

342    rain cells predicted in NU-WRF.

---

[3] http://iphex.pratt.duke.edu/

343    The hindcast results (here only results using Q3/MRMS are shown due to similarity with

344    results using Stage IV) show generally improved performance compared to forecasts for most of

345    the basins except Basin 10 (Fig. 5)  and two small headwater catchments in the Upper Yadkin

346    (i.e. Basins11 and 12, not shown) for the May 15 event. The good forecast performance on May

347    15 in Basin10 demonstrates the importance of the accuracy of the QPF forcing: given high

348    quality QPFs, the hydrologic forecasts using the uncalibrated DCHM are very good such as on

349    May 15;  by contrast, note the false alarm on June 13 in the same basin given overestimated

350    QPFs compared with observations on June 12.

351    It should be stressed that the initial streamflow in each basin for the current day forecast

352    was simply based on the discharge observation at the basin outlet at the time of forecast, i.e.

353    discharge observations were nudged into the DCHM for each basin outlet and proportionally

354    estimated flow redistributed through the basin's channel network according to the ratio of

355    predicted to the observed streamflow at the basin outlet (as described earlier, see workflow in

356    Fig. 2). However, nudging discharge observations at the basin outlet directly into the model

357    could only affect the model states directly tied to river water stage and for a certain (short) period

358    of time as antecedent soil moisture conditions control rainfall-runoff response, as illustrated by

359    the shift in the streamflow curve at the beginning of each day in Figure 4.   This problem can be

360    alleviated   by   assimilating   discharge   observations   into   the   DCHM   to   systematically

361    update/improve soil moisture within the basin.  This is further discussed in section 4.3.

362

363    **3.2 Case study with multiple QPEs**

364    The largest region-wide rainfall event on May 15 with large streamflow response in all 12

365    basins during the IPHEx-IOP is examined closely. A second event, a localized rainfall event on

17

366     June 12 which caused streamflow response in Basin 2 next day (June 13), is not shown here.

367     Figure 5 shows daily rainfall accumulations on May 15 from multiple QPEs (including Stage IV,

368     Q3 and also IMERG) and QPFs from NU-WRF. It can be seen from the figure that Stage IV and

369     Q3 show very similar storm patterns although Q3 patters exhibit sharper spatial variability due to

370     higher resolution. The IMERG data exhibit spatial variability consistent with Stage IV and Q3 at

371     coarse resolution (~10km; e.g. Nogueira and Barros, 2015), but much heavier rainfall for the

372     event in question. That is, the overestimation is preserved by the downscaled product. Moreover,

373     the spatial patterns of NU-WRF QPF do not agree with the QPEs with much larger rainfall

374     accumulations compared to Stage IV and Q3, thus causing significant streamflow overestimation

375     as pointed out earlier. Hindcast results using Stage IV are larger than those using Q3 except for

376     Basins 3 and 5, where both products are similar (Figure 5).  This is illustrated in Figure 6 which

377     exclusively shows daily simulation results for May 15, including hindcasts driven by both

378     StageIV and Q3, as well as the forecasts initialized using the two hindcasts.  The initial

379     conditions for the forecasts or the final states between the two hindcasts for the previous day are

380     very close, consequently leading to very similar performance except for Basin 1. The similarity

381     is explained by the antecedent conditions, specifically a dry period of about two weeks with little

382     antecedent rainfall as indicated in Figure 4, during which the evolution of soil moisture states

383     was controlled by evapotranspiration and deep percolation, and thus antecedent conditions were

384     not affected by Stage IV or Q3. The exception in Basin 1 is caused by discrepancy of rainfall on

385     May 13 between Stage IV and Q3 (not shown here), which leads to large differences in initial

386     conditions for the May 15 event forecast.

387         Figure 7 shows the rainfall accumulation on May 15 from NOXP with two elevation

388     angles at 1.8$^\circ$ and 2.4$^\circ$, and the hybrid data obtained by merging quality observation from various

389    elevation angles. Even though the NOXP was installed at high elevation (as shown in Figure 1)

390    to minimize topographic blocking, the impact of the typical challenges of ground-based radar

391    sensing in mountainous regions, including overshooting, blockage and ground clutter, are

392    apparent in Fig. 8. An overview of hindcast results in the Pigeon River Basin on May 15 using

393    the NOXP data, as well as the NU-WRF QPF and other ground radar-based QPEs including

394    StageIV and Q3, and satellite-based IMERG data, are presented in Fig. 9. Both IMERG and NU-

395    WRF overestimate the rainfall on May 15, thus leading to larger streamflow response.

396    Simulations forced by NOXP QPEs largely underestimate streamflow for all the three small

397    basins in the Pigeon (Basins 1, 2 and 3) due to terrain blocking as stated earlier.

398        A posteriori analysis of hydrologic forecasts and hindcasts indicates that, despite the

399    unusual high density and unique combination of IPHEx observations in this region, "true"

400    rainfall during the IOP remains elusive at this time, though ongoing and future studies will

401    reduce uncertainty through physically-based comprehensive integration of the full suite of

402    IPHEx observations not yet available (Barros et al. 2014).  However, with multiple QPEs and

403    QPFs in hand, a distribution of streamflow simulations can be assembled, the spread of which

404    explicitly represents the propagation of rainfall uncertainty to the hydrologic forecast, or in other

405    words the model's sensitivity to rainfall uncertainty which is essential for quantifying the

406    probability of flood occurrence.  A significant effort was devoted to explore alternative strategies

407    to improve the flood forecasts and hindcasts in the post-IOP phase of IPHEx including better

408    QPF and QPE accuracy, and assimilation of discharge at the forecast points.

409

19

## 4. Improving Results – Meet the challenge

### 4.1 Improving forecasts by enhancing QPFs

The NU-WRF ensemble data assimilation system was developed with a focus on assimilating satellite precipitation-affected radiances into NU-WRF. The system uses an all-sky radiative transfer algorithm to connect the observed microwave radiances with the forecast model states. The analysis control variables are wind, temperature, surface pressure, water vapor and five hydrometeors including frozen and liquid phases. An ensemble of NU-WRF model forecasts are used to calculate state-dependent background error covariance (Zhang et al., 2013; Zupanski et al., 2011). The GPM (Global Precipitation Measurement satellite mission, Matsui et al., 2013) core observatory launched in February 2014 has an orbit extended to higher latitudes (65°) to provide broader spatial coverage (Hou et al. 2014). The microwave imager on board GPM (GMI, Global Microwave Imager) has thirteen microwave channels ranging in frequency from 10 GHz to 183 GHz. There were two overpasses of the GPM core observatory during the May 15 event, providing passive microwave observations of the storm precipitation process from space. To take advantage of these two overpasses, a data assimilation experiment was conducted to assimilate GPM data into NU-WRF, specifically GPM core and constellation cross-calibrated level-1C data from GMI and SSMIS (Special Sensor Microwave Imager/Sounder), aiming at improving the NU-WRF QPF.

The experiment consists of 32 ensemble forecasts and the assimilation cycling is initiated by GFS (Global Forecast System, http://www.emc.ncep.noaa.gov) global analysis at 15UTC May 14, 2014. The assimilation time window is 3 hours. Observations that are available in each assimilation time window are submitted to pass quality control, and a subset of the data are used in the analysis. Two runs were carried out for the cycling period from 15UTC May 14 to 00UTC

433    May 16, 2014. The first run assimilates ground-based conventional data from the NCEP

434    (National Center for Environmental Prediction) data stream including wind, temperature and

435    moisture (denoted as DA-CNT). The second run assimilates GMI and SSMIS (Special Sensor

436    Microwave Imager/Sounder) precipitation-affected microwave radiances at frequencies 89, 166

437    and 183+/-7 GHz (denoted as DA-SAT). The analysis is solved in the outer domain at 9km

438    resolution, and results are dynamically downscaled to 1km resolution via model simulations in

439    the inner domain. Because of prohibitive high computational expense of using large high-

440    resolution domains in ensemble data assimilation cycling, the areal extent of the model domain

441    configuration in these runs is about half of the size of the NU-WRF operational forecast run

442    depicted in Figure 2, and with 31 vertical levels instead of 61 to strike a balance between

443    desirable domain size and vertical resolution and computational costs. The Goddard 3ICE

444    microphysics scheme is applied in model state propagation and in precipitation-affected radiance

445    simulation.

446        The daily accumulations of QPFs from the two assimilation experiments on May 15,

447    2014 are displayed in Figure 5. Comparing to Q3 data and the operational NU-WRF forecast, the

448    storm front traveled rapidly eastward in the control run DA-CNT, resulting in a significant

449    displacement of the spatial QPF pattern. The assimilation run DA-SAT shows improved spatial

450    rainfall patterns and position relative to the control run, but fails to correct the storm cumulative

451    precipitation. The heaviest rain cell is much closer to the actual location as shown in Q3, though

452    with slightly deviated position, i.e. the Q3 displays the heaviest rainfall over the southeast ridge

453    lines of the Upper French Broad River basin, while the heaviest rain cell in the NU-WRF QPF

454    with DA-SAT is on the west ridge lines reaching into the Pigeon River Basin. The flood

455    forecasting results using the two QPFs are provided in Figure 9. Comparing to the streamflow

456    observations and operational forecast-driven results, the QPFs from DA-SAT lead to excessively

457    high streamflow response in the three small headwater catchments of the Pigeon River (Basins 1,

458    2 and 3), while the QPFs from DA-CNT generate much lower streamflow response in the two

459    basins on the eastern slopes of the Appalachians (Basins 2 and 3).  In the inner mountain region,

460    where orographic modulation of precipitation takes place at the ridge-valley scale, the QPFs are

461    too high thus leading to excessive streamflow in Basin 1. These results show that despite clear

462    improvement of the NU-WRF storm forecast with the assimilation of satellite data correcting the

463    storm path and the overall spatial pattern of precipitation as shown by the difference between the

464    accumulated QPFs of DA-CNT and DA-SAT, the improvement takes place at the mesoscale, and

465    thus it's not sufficient to improve the QPF at the headwater catchment scale.  This calls for

466    investigating further refinements in the dynamical downscaling design NU-WRF model

467    configuration and spin-up, and error characterization (e.g. bias) in the radiance assimilation

468    scheme. In this case, the streamflow observations provided valuable verification for satellite data

469    assimilation in hydrological applications, which can serve as a reference point to improve the

470    bias correction in assimilation algorithms and ensemble forecasts.  Finally, because the DA of

471    microwave radiances introduced such a dramatic correction on the position and pattern of the

472    storm, there is also an opportunity to investigate physical-statistical downscaling approaches

473    (e.g. Nogueira and Barros, 2014b)  to leverage the benefits at the mesoscale by improving the

474    representation of moist processes at the cloud-resolving scale that is critical to resolve the

475    individual storm cells that determine streamflow (and flash-flood) response   in mountainous

476    regions.

477

478    **4.2 Improving hindcasts by enhancing QPEs**

479        Previous work has demonstrated success using raingauge observations to characterize

480    errors and uncertainties in QPEs, and then to adjust the QPEs leading to significant

481    improvements in streamflow simulations (Tao and Barros, 2014a; Tao and Barros, 2013).  The

482    same approach was followed to improve the Q3 data. Specifically, the Q3 data were first

483    compared against rainfall observations from the dense raingauge network comprising NASA

484    dual-platform gauges, Duke PMM gauges, HADS and ECONet gauges as shown in Figure 1c,

485    and then were adjusted at hourly time steps by linear regression between the Q3 and gauge

486    observations. Figure 10 shows the comparison between the rainfall observations and the Q3 data,

487    as well as the adjusted Q3 data (noted as Q3+) by three adjusting methods, namely Q3+_All

488    based on the linear regression model derived using all the raingauge observations, Q3+_H/L

489    separating adjustments for high elevation from low elevation as described in Tao and Barros

490    (2013), and Q3+_CdfThr separating heavy rainfall domain from non-heavy rainfall domain using

491    a threshold at 0.9 CDF (cumulative distribution function) derived from raingauge observations

492    (Lin et al., 2015). As it can be seen from the figure, the accuracy of Q3+ is improved with

493    reduced RMSE compared to the original Q3 data, and with relative larger storm rainfall

494    accumulations although differences among the three gauge-corrected Q3+ data sets are small.

495    The adjustments also include value-added information on spatial variability as illustrated by the

496    contrasts between the cumulative rainfall patterns from the original Q3 and the Q3+ data on May

497    15 (Figure 11).  Basin 2 streamflow hindcasts using Q3+ are higher and in  better aggrement with

498    observations, but streamflow is overestimated in Basins 1 and 3 (Figure 12). This highlights the

499    difficulty in capturing small-scale precipitation variability using empirical (data-driven)

500    raingauge correction methods. The number and distribution of gauges is limited in Basin 3 due to

501   the fact that it was not possible to obtain gauge installation permits in the Pisgah National Forest.

502   Moreover, in retrospect, the number of raingauges at mid and low elevations in Basin 1 is

503   insufficient reflecting low awareness of the dominant role of low level orographic rainfall

504   enhancement processes such as seeder-feeder interactions (Wilson and Barros, 2014; Wilson and

505   Barros, 2015) in the design of the raingauge network at the time (2007) when it was first

506   deployed (Prat and Barros, 2010). Consequently, the complexity of orographic modulation of

507   precipitation processes in the SAM is not fully captured at the ridge-valley scale.

508       One of the merits of the simple linear regression adjustment is that the uncertainty

509   associated with Q3 data can be explicitly represented for each pixel at each time step assuming

510   that the uncertainty is normally distributed with the mean as the 'optimum' Q3+ data and

511   standard deviation based on a selected confidence interval (CI) of the regression model, hence

512   providing an unambiguous straightforward framework to specify temporal and spatial error

513   structures in rainfall. The grey lines in Figure 12 depict the streamflow hindcasts spread for 50

514   rainfall replicates drawn from the normal distribution within 70%CI and 95%CI based on the

515   derived regression models for Q3+_All as an example. Note that, even though the QPF from

516   NU-WRF substantially overestimates rainfall, the estimated streamflow is still within the 95%CI

517   envelope, but outside or at the edge of the 70%CI envelope, except for the flow peaks. This

518   implies that all the uncertainty and errors associated with (and not only in) rainfall forcing, but

519   also in initial conditions, model structure and model parameters interact nonlinearly and are

520   propagated and integrated over time leading to the large bias in simulation results. To counteract

521   the compounded effect of error propagataion and model memory on uncertainty build-up,

522   physically-based merging of discharge observations and model forecasts is explored next using

523   data-assimilation techniques.

524

**4.3 Improving forecast/hindcast by assimilating discharge observations**

*4.3.1 Implementation*

525

526

527        To investigate the value of data assimilation (DA) in aiding operational flood forecasts,

528    discharge observations at the basin outlet are assimilated into the DCHM to systematically

529    reduce uncertainty and errors in estimated soil moisture within the basin and thus produce better

530    initial conditions for streamflow forecasting generally and flood forecasting in particular. Three

531    data-assimilation systems (DAS, see the Appendix for detailed mathematical formulation),

532    specifically the Ensemble Kalman Filter (EnKF), the fixed-lag Ensemble Kalman Smoother

533    (EnKS) and the Asynchronous Ensemble Kalman Filter (AEnKF) are tested here. Two models

534    are involved in data assimilation, including a state equation or an input-to-state forward model

535    which propagates hydrological states in time (i.e. the Eq. (1) in the Appendix), and a state-to-

536    output observations operator that relates states to observations (i.e. the Eq. (2) in the Appendix).

537    In this study, the state vector consists of control variables including soil moisture from the top

538    three model soil layers (top, middle and deep layer) at all pixels within the basin. The assimilated

539    observations are the discharge at basin outlets when they become available.  Furthermore, to

540    evaluate a broad range of potential operational data-assimilation architectures, the DAS are

541    implemented in different configurations with regard to assimilation frequency (AF: 15, 30 and 60

542    minutes) and assimilation time window (TW: 1, 2, and 3 hours), as summarized in Table 2.  In

543    the EnKF and EnKS DAS, only the current discharge observations are assimilated, while in the

544    AEnKF all the available discharge observations within the TW are assimilated.

545        When assimilating discharge into a distributed hydrologic model that simulates the space-

546    time evolution of rainfall-runoff response processes, there is a time-lag between the basin

547    internal states at local places (i.e. soil moisture) and the discharge at the basin outlet reflecting

548    the trajectory and travel time of a control volume of runoff (surface or subsurface) from any

549    generic location within the basin to the outlet. The EnKF assimilates the current observation to

550    correct/update the current hydrological states; thus, it does not account for the response delay at

551    the outlet.  The AEnKF is equivalent to a 4D-Var (Four-Dimensional Variational) method but

552    does not need a tangent linear or adjoint model (Sakov et al., 2010), and it accounts for

553    discrepancies among past model predictions and observations also at times different from the

554    assimilation time within the specified TW. The EnKS implemented in this work uses the current

555    observations to correct the antecedent states in the past, propagating information back in time

556    and space to account for the time-lag explicitly, thus effectively re-initializing the model to

557    propagate the updated past states forward to current time. Both the EnKS and AEnKF are

558    asynchronous KF-based (Kalman Filter) algorithms with documented success in improving the

559    representation of the impact of the time-lag in rainfall-runoff response at the outlet on

560    streamflow simulations (Li et al., 2015; Li et al., 2013; Li et al., 2014; Rakovec et al., 2015;

561    Sakov et al., 2010).

562        To generate the model ensembles, stochastic perturbations were applied to atmospheric

563    forcing fields provided by NU-WRF, soil parameters and discharge observations in order to

564    account for associated uncertainties in model inputs and possible measurement errors. Soil

565    moisture estimates were also directly perturbed to account for potential errors in the state

566    forecast model. Table 3 summarizes the methods and parameters applied for each perturbation.

567    QPFs were perturbed by multiplying a realization drawn from a log-normal distribution. Log-

568    normally distributed multiplicative perturbations were also applied to incoming shortwave

569    radiation, while normally distributed additive perturbations were applied for other atmospheric

570    forcing fields including incoming longwave radiation, air temperature, air pressure, specific

571    humidity and wind speed. Soil parameters used for calculation of  the unsaturated hydraulic

572    conductivity ($\mathbf{K}(\boldsymbol{\theta}) = \mathbf{K_s} \left(\frac{\theta}{\emptyset}\right)^{\mathbf{n}}$)(Campbell, 1974), including the saturated hydraulic conductivity

573    $\mathbf{K_s}$ and the power n=3+2/$\lambda$ in which $\lambda$ is the pore-size index, were perturbed using the normally

574    distributed additive method also. The perturbation to static soil parameters is applied once before

575    running the simulations. Spatial soil moisture perturbations were generated by adding normally

576    distributed noise with zero mean and a standard deviation as 5% of top soil moisture at each time

577    step (i.e. 5min). At each location, the spatial soil moisture perturbations were transferred to the

578    top, middle and deep soil layers using relative weights 4:2:1 in an attempt to capture the

579    differences in DCHM soil layer depth and soil hydraulic properties. For the discharge

580    observations, the normally distributed additive perturbation was used with a time-varying

581    standard deviation that is a function of discharge itself, assuming that the uncertainty in

582    discharge  is much larger at high river-stage levels than at low stage levels (Clark et al., 2008;

583    Sorooshian and Dracup, 1980).  Landscape properties such as land-cover, emissivity, albedo, etc.,

584    were not perturbed. Finally, hindcasts were simulated using the Q3+_All gauge-corrected QPE

585    product with uncertainty identified within 95% CI of the adjusting linear regression model as

586    described in section 4.2.

587        The workflow of discharge assimilation is mapped in Figure 13. The latency of discharge

588    observations is 30min~1hour, while the total number of discharge observations assimilated into

589    the DCHM depends on the assimilation frequency, and also the time window for the AEnKF

590    (Table 2). Given the uncertainty described above, a number of replicates of the state vector are

591    propagated in time by the DCHM. At DA time, the true state vector conditioned on observations

592    can be obtained by updating each replicate (background estimate) using a Kalman Gain (KG)

593    matrix $K(t) = C_{XM}(C_M + C_Z)^{-1}$ where $C_{XM}$ is the error cross covariance between state vector

594    and estimated measurements, and $C_M$ and $C_Z$ is are error covariance matrices associated with the

595    predicted measurements (i.e. streamflow estimates) and the observations, respectively. The

596    calculation of KG is different for each tested DA scheme, i.e. AEnKF calculates the KG by

597    augmenting the state vector with past streamflow estimates, while the soil moisture in the

598    calculation of KG for EnKS is at a past time determined by the TW and AF (see details in the

599    Appendix). EnKS is able to update all states within a TW, but here only the first states within the

600    TW (i.e. at $t - TW$) are updated, and next the DCHM propagates the past states from all

601    ensemble members at $(t - TW)$ to the current time $(t)$ again. The process is repeated iteratively

602    at the next assimilation time (as shown in the Figure 13).

603    *4.3.2 Analysis of DAS Performance*

604         Assimilation experiments were conducted in the three basins in the Pigeon River Basin

605    (Basins 1, 2 and 3) for the largest event during the IPHEx-IOP (May 15) only due to the

606    availability of Q3+_All rainfall (refer to Section 4.2). Hindcast results are shown in Figure 14,

607    organized in four panels to illustrate hindcast results for the various DAS configurations: a) using

608    the EnKF with different AF, b) using the AEnKF with different AF and TW, c) using the EnKS

609    with different AF and TW, and d) the three best DAS identified according to the NSE (Nash-

610    Sutcliffe Efficiency) metric as summarized in Table 2. Other evaluation metrics including the

611    KGE (Kling-Gupta Efficiency) and the modified KGE (Gupta et al., 2009; Kling et al., 2012),

612    and the errors in the peak flow value (EPV) and time (EPT) are also provided. It can be seen

613    from Figure 14 that the EnKF is not capable of correctly capturing the temporal lag between

614    basin states and basin-output fluxes during rainfall, because updating soil moisture storage at the

615    DA time corrects the current discharge but it does not account for the time delay required to

616    transfer the joint effects of spatial variability of antecedent soil moisture and rainfall on runoff

617    generation to the basin outlet. By contrast, by also assimilating past discharge observations, the

618    AEnKF produces much better simulations especially in Basins 1 and 3 compared to EnKF. The

619    simulations with AEnKF are particularly improved for Basin 3 (AF = 15min; TW = 2hrs) with

620    the NSE, KGE and modified KGE equal to 0.99, 0.94 and 0.96, respectively. The EnKS DAS

621    also show better performance than EnKF due to explicitly accounting for the time-lag between

622    basin internal states and outlet response, attaining an NSE, KGE and modified KGE of 0.98, 0.95

623    and 0.97 for Basin 1 (AF = 15 min; TW = 2 hrs). Note that, as pointed out by Tao and Barros

624    (2013), both Basin 1 and Basin 3 have deep alluvial valleys which naturally slow and smooth

625    rainfall-runoff response, and thus the hydrological processes are amenable to time integration at

626    moderate temporal resolution. The nearly perfect skill achieved for AEnKF and EnKS

627    configurations is partly attributed to the AF, i.e. the best performance is achieved by

628    assimilating as many discharge observations as possible, and thus the optimal AF is equal to the

629    discharge observation frequency (every 15min) consistent with Wanders et al. (2014). A note of

630    caution is warranted as KF-based DAS implementations imply that observation errors are serially

631    independent, an assumption that can be compromised when streamflow observations are very

632    close together in time. However, given the large background uncertainty as shown in the Figure

633    14d) and the small uncertainty associated with observations (std. specified as 10% of the

634    observations), this is it not likely to be a significant issue for this particular assimilation problem.

635    Finally, AEnKF displays relatively lower uncertainty (shown by the ensemble spread for Basin 3

636    in Figure 14d) than EnKS (shown by the ensemble spread for Basins 1 and 2 in Figure 14d) by

637    assimilating many (past) discharge observations, not just the current one.

638    Nevertheless, none of the DAS shows good results for Basin 2, the smallest basin with

639    drainage area of 71km$^2$, steep slopes and shallow soils. The Basin 2 simulation with a best NSE

640    of 0.71 is produced by EnKS with 15min AF and 1hr TW. Although the major peak of the

641    hydrograph is underestimated and the KGE and the modified KGE are relatively low (0.58 and

642    0.72, respectively), the peak time error is among the smallest (±30min), which is critical for

643    flash-flood warning, and thus we still use this scheme (AF = 15min; TW = 1hr) as the best

644    configuration for Basin 2. Simulations with longer TW, i.e. EnKS_AF15min_TW2hr and

645    EnKS_AF15min_TW3hr, show comparable or slightly worse NSE results (0.67 and 0.61,

646    respectively as shown in Table 2) but have significant better KGE, modified KGE and peak

647    values, albeit with larger errors in time-to-peak (about 1.5 hr). That is, the EnKS updating of

648    antecedent soil moisture 2hr or 3hr before the assimilation time has a strong impact on the

649    streamflow at the basin outlet 0.5-1.5 hr later, thus over a shorter time-lag than the TW (2-3hr).

650    This behavior implies that the weights used to transfer soil moisture perturbations in the different

651    soil layers are important to determine the simulated hydrograph ensemble spread when the

652    number of ensemble replicates is limited. For example, surface runoff and shallow interflow

653    dominate the rising limb of the hydrograph in Basin 2 (Barros and Tao, 2013) and therefore the

654    amplitude of soil moisture perturbations in the two top soil layers will determine the spread of

655    the simulated discharge in this case. Understanding of rainfall-runoff processes in the context of

656    basin-specific topography and geomorphology can provide therefore valuable insights in the

657    practical implementation of ensemble-based DAS.

658    Previous studies suggest that the time of concentration is a good estimate of the TW for

659    DA (Li et al., 2013; Rakovec et al., 2015). However, the experiments conducted in the context of

660    this work suggest that quality DAS is associated with TWs significantly shorter than the time of

661    concentration (e.g. about 5hr for the smallest Basin 2, and much larger for Basins 1 and 3).

662    Indeed, the best performance is attained when the latency of the observations is assumed to be

663    nearly instantaneous (AF=temporal resolution of the observations), which is possible for these

664    hindcast simulations, but unrealistic in an operational environment.  It should be emphasized that

665    for distributed hydrologic models the DAS performance for a particular basin depends not only

666    on basin geomorphologic features (i.e. topography, elevation, size, etc.) but also on temporal and

667    spatial rainfall characteristics (i.e. rain cell's location is close to the basin outlet or not), initial

668    soil moisture conditions, and their uncertainty.   Although there is no universal DAS

669    configuration that will outperform all others at all times, a priori studies to explore the sensitivity

670    of DAS to the TW/AF ratio that is ultimately controlled by the temporal resolution of the

671    observations and their latency should prove helpful in practice.

672    *4.3.3 Operational Forecasting Application*

673         Here, we use the 'best' DAS from the flood hindcast simulations for each basin (i.e.

674    EnKS_AF15min_TW2hr  for  Basin  1,  EnKS_AF15min_TW1hr  for  Basin  2  and

675    AEnKF_AF15min_TW2hr for Basin 3, Table 2) to simulate flood forecasting in operational

676    mode, i.e. assimilating available discharge observations only before the forecasting time

677    (illustrated by Figure 13).

678         The flood forecasting results assimilating discharge observations are presented in Figure

679    15, and the corresponding evaluation metrics are summarized in Table 4.  As discussed earlier,

680    the purpose of asynchronous and smoother implementations of the Kalman Filter is to introduce

681    memory in the data assimilation and thus capture nonlinear interactions that are essential to

682    improve initial conditions for future forecasts.   This is apparent from inspecting the EnKS

683    results:  the soil moisture storage at $t$-TW is improved by assimilating observations at time $t$, and

684     the updated states at $t$-TW (i.e. improved initial conditions for $t$-TW+1 onward), were propagated

685     subsequently by the DCHM to time $t$. From the point of view of capturing the highly-nonlinear

686     rainfall-runoff processes, the states propagated to $t$ after correction by the EnKS at $t$-TW are

687     more accurate than the original states at $t$, or the updated states at t by EnKF (i.e. improved initial

688     conditions for $t$ +1 onward, which is to say the EnKS updating at $t$-TW is equivalent to model re-

689     initialization). In the context of operational forecasts, the maximum forecast lead time is the time

690     difference between the last step of the forecasting simulation (00UTC) and the time when the

691     forecast is issued (as indicated by the dots on the time-axis in Figure 15). For Basins 2 and 3, the

692     forecasting results with shorter lead times are better than with longer lead times as expected

693     (NSEs are summarized in Table 4). Interestingly, for Basin 1, forecast skill is best for the 12hr-

694     lead time.   This behavior is explained by the temporal variability of rainfall over the basin: the

695     predicted storm (QPF) began around 03UTC for all three basins, and it lasted until 11UTC in

696     Basins 2 and 3 but it stopped sharply before 09UTC in Basin 1, thus explaining the maximum

697     lead time of 15 hours.  Assimilating discharge after the storm stops does not add forecast value

698     because the uncertainty in rainfall is specified as a fraction of the QPF, and the corrections

699     applied to the model state vector are too small despite large streamflow innovations. In Basins 2

700     and 3, the major storm activity stopped around 07UTC, but it was followed by two smaller

701     events that are essential to widen the ensemble spread of the simulations, and thus enable

702     discharge assimilation to add information (i.e. observations are within the estimation space).

703     Exploring strategies to represent uncertainty in the timing of rainfall onset and termination,

704     conditional on local hydrometeorology and specific storm characteristics, should help with

705     improving DAS performance, especially in small basins and for short heavy precipitation events

706     which are critical for flash-flood forecasting.  Finally, note very large NSEs of 0.87, 0.78, 0.72

707    and 0.51 for flood forecasting in Basin 3 for 6hr, 9hr, 12hr and 15hr lead times, a robust

708    performance that is  uncommon in operational flood forecasting, especially using uncalibrated

709    physically-based hydrologic models (e.g. Kim and Barros, 2001 for results using data driven

710    models).

711

712    **5.   Conclusions and Discussion**

713         During the IPHEx-IOP, daily flood hindcasts and forecasts were conducted in a virtual

714    operational environment without tuning initial conditions or model calibration for twelve

715    headwater catchments in the Southern Appalachians. In the post-IOP phase of the campaign,

716    various strategies were implemented in order to investigate alternative pathways to improve

717    flood forecasting skill in mountainous regions including: improvement of NWP QPFs,

718    improvement of QPEs with an eye on improving initial conditions for hydrologic modeling, and

719    improvement of QFFs (Quantitative flash-Flood Forecasts) through assimilation of discharge

720    observations.  The latter proved to be the most promising approach attaining superior (an

721    unprecedented) skill for long lead-times in headwater basins. The study also illustrated the

722    sensitivity of DAS to basin hydro-geomorphic characteristics in addition to the temporal and

723    spatial structure of rainfall: a survey of Table 2 shows that DCHM-DAS skill metrics for Basins

724    1 and 3, larger watersheds with alluvial valleys and slower rainfall-runoff response, are

725    significantly less variable among the various configurations than the skill metrics for Basin 2, a

726    small catchment with shallow gravelly soils and steep slopes.

727         Future operational testbeds could benefit from multi-model QPFs and multi-model QFFs

728    (i.e. using multiple hydrological models with multi-source of QPFs to produce a multi-model

729    streamflow ensemble), implementation of operational forecasting with longer lead times on the

730    basis of local time (instead of UTC time), near-real time ingestion of ground- and satellite-based

731    QPEs, and assimilating not only discharge observations, but also satellite-based and/or ground-

732    based soil moisture observations, to improve initial for hydrological forecasts. The latter can

733    provide valuable constraints to address the question of uncertainty in the choice of the

734    assimilation time window as the antecedent space-time variability of rainfall can be characterized

735    by the soil moisture products, i.e. estimating a suitable time window based on temporal-spatial

736    soil moisture information for each assimilation time.   Specific opportunities for improving a

737    number of issues are worthwhile further investigation:

738         i) The discharge assimilation show significant flood forecasting improvements for

739    individual events during the IPHEx-IOP.   During wet periods, the benefits of continuous DAS,

740    specifically by correcting soil moisture, may lead to even better results by providing better initial

741    conditions for sequential storms. Nevertheless, only one major storm occurred during the IPHEx-

742    IOP, and further evaluation of the coupled DCHM-DAS should be pursued for a larger number

743    of storms encompassing representative synoptic and mesoscale weather regimes.. This could be

744    accomplished in the future by selecting a historical period with  several successive events for

745    investigating of the system's effectiveness in improving initial conditions of later events by

746    assimilating discharge observations of preceding events. Further work is also needed to

747    implement the data assimilation systems tested here in realistic operational environments.

748         ii) Even though a unique combination of high-quality QPE products was obtained for the

749    campaign, none of these are perfect, i.e. raingauge data only represent point-scale observations,

750    ground-based radar observations severely suffer from topography related errors in mountainous

751    regions, and satellite-based observations are limited by retrieval uncertainty and typically have

752  coarse spatial and, or temporal resolution.   Assimilating discharge data for correcting rainfall

753  and model parameters using lumped hydrologic models was pursued previously (Harader et al.,

754  2012), but it had not yet been attempted using a fully-distributed model in mountainous terrain.

755  Further research is needed to integrate the benefits of improved QPFs and QPEs with hydrologic

756  DAS.

757          iii) Because landslides (e.g. debris flow) are linked often to flood events in mountainous

758  terrain, there is an opportunity to further extend the operational flood forecasting framework to

759  include landslide initiation as in Tao and Barros (2014a).

760

## Appendix: Data Assimilation Algorithms

Data assimilation schemes include two models, a state equation or an input-to-state forward model (the physics model) that propagates hydrologic states in time, and an observation operator or a state-to-output model that relates hydrologic states with observations (Liu and Gupta, 2007). The forward model is represented using Equation (1),

$$x(t) = \mathcal{F}(x(t-1), \alpha, u(t), t) + \omega(t) \tag{1}$$

where x(t) is the state vector, $\mathcal{F}$ is the DCHM in our case, $\alpha$ represents time-invariant data sets or model parameters, u(t) represents time-variant forcing data sets, and $\omega(t)$ is the uncertainty in the model structure. Given appropriate uncertainty representation, an ensemble of a number of replicates of the state vector is propagated from t-1 to t. Each replicate of the state vector can be written as $x_j(t)$ where j is the $j^{th}$ replicate of an ensemble of size **Ne**. In this study, the control variables include soil moisture from each soil layer at all the pixels within a basin, i.e. $x_j = [\theta_1^t, ..., \theta_N^t, \theta_1^m, ..., \theta_N^m, \theta_1^d, ..., \theta_N^d]_j^T$ where $\theta_*^t$ is the soil moisture in the top soil layer, $\theta_*^m$ is the soil moisture in the middle soil layer, and $\theta_*^d$ is the soil moisture in the deep soil layer. **N** is the total number of basin grid elements. The size of the state vector $x_j$ is Ns × 1, where **Ns** (**Ns** = 3N) is the total number of control variables or states.

The observations operator $\mathcal{M}$ maps the true state vector to the observations vector $z(t)$,

$$z(t) = \mathcal{M}(x(t^*)) + \xi(t) \tag{2}$$

where $\xi(t)$ represents the uncertainty associated with the observations, distributed with a zero mean and a covariance matrix $C_Z$. Here $z(t)$ are the discharge observations at basin outlets, and

792   thus $\mathcal{M}$ represents the non-linear hydrological processes converting soil moisture states to the

793   basin discharge, which indeed is a Markov process relating observations not only to the states at

794   current time but also at antecedent time steps (indicated by $t^*$). The various ensemble data

795   assimilation schemes differ in the updating strategies.

796        **a) Ensemble Kalman Filter (EnKF) and Asynchronous EnKF**

797        In the EnKF, the updating equation is given by,

$$x_j^+(t) = x_j(t) + K(t)\left(z_j(t) - \mathrm{M}\left(x_j(t)\right)\right) \tag{3}$$

798   where $x_j^+(t)$ represents the updated states (posterior or analysis) and $x_j(t)$ is the state vector

799   before updating (prior or background estimates), $\mathrm{M}\left(x_j(t)\right)$ is the j$^{th}$ replicate of streamflow

800   estimates by the DCHM, and $K(t)$ is the Kalman gain matrix calculated as follows:

$$K(t) = C_{XM}(C_M + C_Z)^{-1} \tag{4}$$

801   $C_{XM}$ is the error cross covariance between state vector and estimated measurements at current

802   (DA) time t, and $C_M$ and $C_Z$ are the error covariance matrices associated with the predicted

803   measurements and the observations, respectively.

804        The Asynchronous EnKF (AEnKF) is a modified version of the EnKF recently proposed

805   by Sakov et al. (2010), which accounts for mismatches between historical estimates and

806   observations at times different from the assimilation time. The updating equation for the AEnKF

807   is expressed by Equation (6),

$$x_j^+(t) = x_j(t) + K_{TW}\left(z_j^T - M_j^T\right) \tag{5}$$

808    where the Kalman gain matrix $K_{TW}$ is calculated by augmenting the state vector with past model

809    predictions within an assimilating time window (TW) (see details in (Rakovec et al., 2015)), and

810    the transpose vectors $z_j^T$ and $M_j^T$ include all the observations and model predictions within the

811    TW. Note that the dimension of $K_{TW}$ is different from $K(t)$ in Equation (4).

812    **b) Ensemble Kalman Smoother (EnKS)**

813    In the EnKS, the updating is not just applied to the current time step, but can be also

814    applied for previous time steps within an assimilating time window (TW). The updating equation

815    of a fixed-lag EnKS is expressed by:

$$x_j^+(t - TW) = x_j(t - TW) + K_{TW}\{z_j(t) - M[x_j(t)]\} \tag{6}$$

816    and the error cross covariance $C_{XM}$ in the Kalman gain matrix $K_{TW}$ is calculated using the

817    antecedent state variables at t-TW and the model predictions at current time t. Others are the

818    same as for equation (3), and the $K_{TW}$ here has the same dimension as $K(t)$ in Equation (4).

819    Equation (6) indicates that the updating procedure can be performed for multiple prior time steps

820    within the TW. However, for physically-based and fully-distributed hydrological models such as

821    the DCHM, the memory of the hydrologic system (e.g. soil water storage in the basin) cannot be

822    directly explained in the EnKS, and thus it needs to be propagated forward by the model itself,

823    that is equivalent to model re-initialization (Li et al., 2015). In this study, only the states at t-TW

824    are updated using Equation (6) and then are propagated in time by the DCHM.

## References

825 **References**

826 Addor, N., Jaun, S., Fundel, F., Zappa, M., 2011. An operational hydrological ensemble

827         prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios.

828         Hydrol Earth Syst Sc, 15(7): 2327-2347.

829 Alfieri, L. et al., 2014. Evaluation of ensemble streamflow predictions in Europe. J. Hydrol., 517:

830         913-922.

831 Amengual, A. et al., 2008. A hydrometeorological model intercomparison as a tool to quantify

832         the forecast uncertainty in a medium size basin. Nat Hazard Earth Sys, 8(4): 819-838.

833 Amengual, A., Romero, R., Vich, M., Alonso, S., 2009. Inclusion of potential vorticity

834         uncertainties into a hydrometeorological forecasting chain: application to a medium size

835         basin of Mediterranean Spain. Hydrol Earth Syst Sc, 13(6): 793-811.

836 Anagnostou, M.N. et al., 2013. Performance Evaluation of a New Dual-Polarization

837         Microphysical Algorithm Based on Long-Term X-Band Radar and Disdrometer

838         Observations. J. Hydrometeorol., 14(2): 560-576.

839 Band, L.E., Hwang, T., Hales, T.C., Vose, J., Ford, C., 2012. Ecosystem processes at the

840         watershed scale: Mapping and modeling ecohydrological controls of landslides.

841         Geomorphology, 137(1): 159-167.

842 Barros, A.P., 1995. Adaptive Multilevel Modeling of Land-Atmosphere Interactions. Journal of

843         Climate, 8(9): 2144-2160.

844 Barros, A.P. et al., 2014. NASA GPM-Ground Validation: Integrated Precipitation and

845         Hydrology Experiment 2014 Science Plan. EPL/Duke University (Pub.): 64.

846    Bartholmes, J.C., Thielen, J., Ramos, M.H., Gentilini, S., 2009. The european flood alert system

847        EFAS - Part 2: Statistical skill assessment of probabilistic and deterministic operational

848        forecasts. Hydrol Earth Syst Sc, 13(2): 141-153.

849    Benoit, R., Kouwen, N., Yu, W., Chamberland, S., Pellerin, P., 2003. Hydrometeorological

850        aspects of the Real-Time Ultrafinescale Forecast Support during the Special Observing

851        Period of the MAP. Hydrol Earth Syst Sc, 7(6): 877-889.

852    Berthet, L., Andreassian, V., Perrin, C., Javelle, P., 2009. How crucial is it to account for the

853        antecedent moisture conditions in flood forecasting? Comparison of event-based and

854        continuous approaches on 178 catchments. Hydrol Earth Syst Sc, 13(6): 819-831.

855    Bloschl, G., Reszler, C., Komma, J., 2008. A spatially distributed flash flood forecasting model.

856        Environmental Modelling & Software, 23(4): 464-478.

857    Campbell, G.S., 1974. A simple method for determining unsaturated conductivity from moisture

858        retention data. Soil Sci., 117(6): 311-314.

859    Casadel, M., Dietrich, W.E., Miller, N.L., 2003. Testing a model for predicting the timing and

860        location of shallow landslide initiation in soil-mantled landscapes. Earth Surface

861        Processes and Landforms, 28(9): 925-950.

862    Castaings, W., Dartus, D., Le Dimet, F.X., Saulnier, G.M., 2009. Sensitivity analysis and

863        parameter estimation for distributed hydrological modeling: potential of variational

864        methods. Hydrol Earth Syst Sc, 13(4): 503-517.

865    Ciach, G.J., Krajewski, W.F., Villarini, G., 2007. Product-error-driven uncertainty model for

866        probabilistic quantitative precipitation estimation with NEXRAD data. J. Hydrometeorol.,

867        8(6): 1325-1347.

868 Clark, M.P. et al., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of
869      streamflow observations to update states in a distributed hydrological model. Adv. Water
870      Resour., 31(10): 1309-1324.

871 Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: A review. J. Hydrol., 375(3-4):
872      613-626.

873 Coccia, G., Todini, E., 2011. Recent developments in predictive uncertainty assessment based on
874      the model conditional processor approach. Hydrol Earth Syst Sc, 15(10): 3253-3274.

875 Collier, C.G., 2007. Flash flood forecasting: What are the limits of predictability? Quarterly
876      Journal of the Royal Meteorological Society, 133(622): 3-23.

877 Connelly, B.A. et al., 1999. Advanced hydrologic prediction system. Journal of Geophysical
878      Research-Atmospheres, 104(D16): 19655-19660.

879 Crow, W.T., Reichle, R.H., 2008. Comparison of adaptive filtering techniques for land surface
880      data assimilation. Water Resources Research, 44(8).

881 Crow, W.T., Van Loon, E., 2006. Impact of incorrect model error assumptions on the sequential
882      assimilation of remotely sensed surface soil moisture. J. Hydrometeorol., 7(3): 421-432.

883 Cuo, L., Pagano, T.C., Wang, Q.J., 2011. A Review of Quantitative Precipitation Forecasts and
884      Their Use in Short- to Medium-Range Streamflow Forecasting. J. Hydrometeorol., 12(5):
885      713-728.

886 Davolio, S. et al., 2009. High resolution simulations of a flash flood near Venice. Nat Hazard
887      Earth Sys, 9(5): 1671-1678.

888 Davolio, S., Miglietta, M.M., Diomede, T., Marsigli, C., Montani, A., 2013. A flood episode in
889      northern Italy: multi-model and single-model mesoscale meteorological ensembles for
890      hydrological predictions. Hydrol Earth Syst Sc, 17(6): 2107-2120.

891   De Jongh, I., Quintelier, E., Cauwenberghs, K., 2012. Using combined raingauge and high-
892        resolution radar data in an operational flood forecast system in Flanders. Weather Radar
893        and Hydrology, 351: 472-477.
894   Demargne, J. et al., 2009. Application of Forecast Verification Science to Operational River
895        Forecasting in the US National Weather Service. Bulletin of the American
896        Meteorological Society, 90(6): 779-784.
897   Demargne, J. et al., 2014. The Science of NOAA's Operational Hydrologic Ensemble Forecast
898        Service. Bulletin of the American Meteorological Society, 95(1): 79-98.
899   Demeritt, D., Nobert, S., Cloke, H.L., Pappenberger, F., 2013. The European Flood Alert System
900        and the communication, perception, and use of ensemble predictions for operational flood
901        risk management. Hydrol Process, 27(1): 147-157.
902   Devonec, E., Barros, A.P., 2002. Exploring the transferability of a land-surface hydrology model.
903        J. Hydrol., 265(1-4): 258-282.
904   Dietrich, J. et al., 2009. Assessing uncertainties in flood forecasts for decision making: prototype
905        of an operational flood management system integrating ensemble predictions. Nat Hazard
906        Earth Sys, 9(4): 1529-1540.
907   Drobinski, P. et al., 2014. HYMEX A 10-Year Multidisciplinary Program on the Mediterranean
908        Water Cycle. Bulletin of the American Meteorological Society, 95(7): 1063-+.
909   Droegemeier, K.K. et al., 2000. Hydrological aspects of weather prediction and flood warnings:
910        Report of the Ninth Prospectus Development Team of the US Weather Research Program.
911        Bulletin of the American Meteorological Society, 81(11): 2665-2680.

912    Ducrocq, V. et al., 2014. HYMEX-SOPI The Field Campaign Dedicated to Heavy Precipitation

913         and Flash Flooding in the Northwestern Mediterranean. Bulletin of the American

914         Meteorological Society, 95(7): 1083-+.

915    Ebert, E.E., 2001. Ability of a poor man's ensemble to predict the probability and distribution of

916         precipitation. Mon. Weather Rev., 129(10): 2461-2480.

917    Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using

918         Monte Carlo methods to forecast error statistics. Journal of Geophysical Research-

919         Oceans, 99(C5): 10143-10162.

920    Evensen, G., 2003. The ensemble Kalman filter: Theoretical formulation and practical

921         implementation. Ocean Dynam, 53(4): 343-367.

922    Evensen, G., van Leeuwen, P.J., 2000. An ensemble Kalman smoother for nonlinear dynamics.

923         Mon. Weather Rev., 128(6): 1852-1867.

924    Ferretti, R. et al., 2014. Overview of the first HyMeX Special Observation Period over Italy:

925         observations and model results. Hydrol Earth Syst Sc, 18(5): 1953-1977.

926    Flores, A.N., Entekhabi, D., Bras, R.L., 2010. Reproducibility of soil moisture ensembles when

927         representing soil parameter uncertainty using a Latin Hypercube-based approach with

928         correlation control. Water Resources Research, 46.

929    Garambois, P.A., Roux, H., Larnier, K., Castaings, W., Dartus, D., 2013. Characterization of

930         process-oriented hydrologic model behavior with temporal sensitivity analysis for flash

931         floods in Mediterranean catchments. Hydrol Earth Syst Sc, 17(6): 2305-2322.

932    Garcia-Pintado, J., Barbera, G.G., Erena, M., Castillo, V.M., 2009. Calibration of structure in a

933         distributed forecasting model for a semiarid flash flood: Dynamic surface storage and

934         channel roughness. J. Hydrol., 377(1-2): 165-184.

935 Garcia-Quijano, J.F., Barros, A.P., 2005. Incorporating canopy physiology into a hydrological

936   model: photosynthesis, dynamic respiration, and stomatal sensitivity. Ecological

937   Modelling, 185(1): 29-49.

938 Gebremichael, M., Barros, A.P., 2006. Evaluation of MODIS gross primary productivity (GPP)

939   in tropical monsoon regions. Remote Sensing of Environment, 100(2): 150-166.

940 Gourley, J.J., Vieux, B.E., 2005. A method for evaluating the accuracy of quantitative

941   precipitation estimates from a hydrologic modeling perspective. J. Hydrometeorol., 6(2):

942   115-133.

943 Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared

944   error and NSE performance criteria: Implications for improving hydrological modelling.

945   J. Hydrol., 377(1-2): 80-91.

946 Hapuarachchi, H.A.P., Wang, Q.J., Pagano, T.C., 2011. A review of advances in flash flood

947   forecasting. Hydrol Process.

948 Hersbach, H., 2000. Decomposition of the continuous ranked probability score for ensemble

949   prediction systems. Weather and Forecasting, 15(5): 559-570.

950 Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., Braatz, D., 2000. A multistep automatic

951   calibration scheme for river forecasting models. J. Hydrometeorol., 1(6): 524-542.

952 Hsiao, L.F. et al., 2013. Ensemble forecasting of typhoon rainfall and floods over a mountainous

953   watershed in Taiwan. J. Hydrol., 506: 55-68.

954 Huffman, G., 2015. GPM Level 3 IMERG Half Hourly 0.1 x 0.1 degree Precipitation,version 03.

955   NASA Goddard Space Flight Center Distributed Active Archive Center (GSFC DAAC),

956   Greenbelt, MD.

957 Huffman, G.J. et al., 2014. NASA Global Precipitation Measurement (GPM) Integrated Multi -
958     satellitE Retrievals for GPM (IMERG), NASA/GSFC, Greenbelt, MD.

959 Jasper, K., Gurtz, J., Herbert, L., 2002. Advanced flood forecasting in Alpine watersheds by
960     coupling meteorological observations and forecasts with a distributed hydrological model.
961     J. Hydrol., 267(1-2): 40-52.

962 Jaun, S., Ahrens, B., 2009. Evaluation of a probabilistic hydrometeorological forecast system.
963     Hydrol Earth Syst Sc, 13(7): 1031-1043.

964 Kalogiros, J. et al., 2013a. Correction of Polarimetric Radar Reflectivity Measurements and
965     Rainfall Estimates for Apparent Vertical Profile in Stratiform Rain. Journal of Applied
966     Meteorology and Climatology, 52(5): 1170-1186.

967 Kalogiros, J. et al., 2013b. Optimum Estimation of Rain Microphysical Parameters From X-Band
968     Dual-Polarization Radar Observables. Ieee Transactions on Geoscience and Remote
969     Sensing, 51(5): 3063-3076.

970 Kalogiros, J. et al., 2014. Evaluation of a New Polarimetric Algorithm for Rain-Path Attenuation
971     Correction of X-Band Radar Observations Against Disdrometer. Ieee Transactions on
972     Geoscience and Remote Sensing, 52(2): 1369-1380.

973 Kang, D., Barros, A., Déry, S., 2013. Evaluating Passive Microwave Radiometry for the
974     Dynamical Transition From Dry to Wet Snowpacks. IEEE Transactions on Geoscience
975     and Remote Sensing.

976 Kang, D.H., Barros, A.P., 2012a. Observing System Simulation of Snow Microwave Emissions
977     Over Data Sparse Regions-Part I: Single Layer Physics. IEEE Transactions on
978     Geoscience and Remote Sensing, 50(5): 1785-1805.

979     Kang, D.H., Barros, A.P., 2012b. Observing System Simulation of Snow Microwave Emissions

980         Over Data Sparse Regions-Part II: Multilayer Physics. IEEE Transactions on Geoscience

981         and Remote Sensing, 50(5): 1806-1820.

982     Kim, G., Barros, A.P., 2001. Quantitative flood forecasting using multisensor data and neural

983         networks. J. Hydrol., 246(1-4): 45-62.

984     Kirstetter, P.-E. et al., 2012. Toward a Framework for Systematic Error Modeling of Spaceborne

985         Precipitation Radar with NOAA/NSSL Ground Radar Based National Mosaic QPE. J.

986         Hydrometeorol., 13(4): 1285-1300.

987     Kling, H., Fuchs, M., Paulin, M., 2012. Runoff conditions in the upper Danube basin under an

988         ensemble of climate change scenarios. J. Hydrol., 424: 264-277.

989     Komma, J., Bloschl, G., Reszler, C., 2008. Soil moisture updating by Ensemble Kalman Filtering

990         in real-time flood forecasting. J. Hydrol., 357(3-4): 228-242.

991     Lee, H., Seo, D.J., Koren, V., 2011. Assimilation of streamflow and in situ soil moisture data

992         into operational distributed hydrologic models: Effects of uncertainties in the data and

993         initial model soil moisture states. Adv. Water Resour., 34(12): 1597-1615.

994     Li, H.B., Luo, L.F., Wood, E.F., Schaake, J., 2009. The role of initial conditions and forcing

995         uncertainties in seasonal hydrologic forecasting. Journal of Geophysical Research-

996         Atmospheres, 114.

997     Li, Y., Ryu, D., Western, A.W., Wang, Q., 2015. Assimilation of stream discharge for flood

998         forecasting: Updating a semidistributed model with an integrated data assimilation

999         scheme. Water Resources Research.

1000    Li, Y., Ryu, D., Western, A.W., Wang, Q.J., 2013. Assimilation of stream discharge for flood

1001        forecasting: The benefits of accounting for routing time lags. Water Resources Research,

1002        49(4): 1887-1900.

1003    Li, Y. et al., 2014. An integrated error parameter estimation and lag-aware data assimilation

1004        scheme for real-time flood forecasting. J. Hydrol., 519: 2722-2736.

1005    Liao, Z. et al., 2011. Evaluation of TRIGRS (transient rainfall infiltration and grid-based regional

1006        slope-stability analysis)'s predictive skill for hurricane-triggered landslides: a case study

1007        in Macon County, North Carolina. Nat Hazards, 58(1): 325-339.

1008    Lin, X., Kidd, C., Tao, J., Barros, A.P., 2015. Comparisons of Rain Estimates from Ground

1009        Radar and Satellite over Mountainous Regions. Journal of Hydrometeorology (In

1010        Revision).

1011    Lin, Y., Mitchell, K., 2005. The NCEP stage II/IV hourly precipitation analyses: Development

1012        and applications, 19th Conf. on Hydrology, , Amer. Meteor. Soc., San Diego, CA.

1013    Liu, Y. et al., 2012. Advancing data assimilation in operational hydrologic forecasting:

1014        progresses, challenges, and emerging opportunities. Hydrol Earth Syst Sc, 16(10): 3863-

1015        3887.

1016    Liu, Y.Q., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data

1017        assimilation framework. Water Resources Research, 43(7).

1018    Lu, C., Yuan, H., Tollerud, E.I., Wang, N., 2010. Scale-Dependent Uncertainties in Global QPFs

1019        and QPEs from NWP Model and Satellite Fields. J. Hydrometeorol., 11(1): 139-155.

1020    Madsen, H., Skotner, C., 2005. Adaptive state updating in real-time river flow forecasting - a

1021        combined filtering and error forecasting procedure. J. Hydrol., 308(1-4): 302-312.

1022    Martina, M.L.V., Todini, E., Libralon, A., 2008. Rainfall Thresholds for Flood Warning Systems:

1023          A Bayesian Decision Approach. Water Trans, 63: 203-227, 291.

1024    Mascaro, G., Vivoni, E.R., Deidda, R., 2010. Implications of Ensemble Quantitative

1025          Precipitation Forecast Errors on Distributed Streamflow Forecasting. J. Hydrometeorol.,

1026          11(1): 69-86.

1027    Matsui, T. et al., 2013. GPM Satellite Simulator over Ground Validation Sites. Bulletin of the

1028          American Meteorological Society, 94(11): 1653-1660.

1029    Matsui, T. et al., 2014. Introducing multisensor satellite radiance-based evaluation for regional

1030          Earth System modeling. Journal of Geophysical Research-Atmospheres, 119(13).

1031    McEnery, J., Ingram, J., Duan, Q.Y., Adams, T., Anderson, L., 2005. NOAA's advanced

1032          hydrologic prediction service - Building pathways for better science in water forecasting.

1033          Bulletin of the American Meteorological Society, 86(3): 375-+.

1034    McLaughlin, D., 2002. An integrated approach to hydrologic data assimilation: interpolation,

1035          smoothing, and filtering. Adv. Water Resour., 25(8-12): 1275-1286.

1036    Moore, R.J., Cole, S.J., Bell, V.A., Jones, D.A., 2006. Issues in flood forecasting: ungauged

1037          basins, extreme floods and uncertainty. Frontiers in Flood Research, 305: 103-122.

1038    Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A

1039          discussion of principles. J. Hydrol., 10(3): 282-290.

1040    Nester, T., Komma, J., Viglione, A., Bloschl, G., 2012. Flood forecast errors and ensemble

1041          spread-A case study. Water Resources Research, 48.

1042    Nogueira, M., Barros, A., 2014. The Integrated Precipitation and Hydrology Experiment -

1043          Hydrologic Applications for the Southeast US (IPHEx-H4SE) Part III: High-Resolution

Ensemble Rainfall Products. Report EPL-2013-IPHEX-H4SE-3, EPL/Duke University (Pub.): 80.

Nogueira, M., Barros, A.P., 2015. Transient Stochastic Downscaling of Quantitative Precipitation Estimates for Hydrological Applications. Journal of Hydrology (In Revision).

Noh, S.J., Rakovec, O., Weerts, A.H., Tachikawa, Y., 2014. On noise specification in data assimilation schemes for improved flood forecasting using distributed hydrological models. J. Hydrol., 519: 2707-2721.

Norbiato, D., Borga, M., Esposti, S.D., Gaume, E., Anquetin, S., 2008. Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins. J. Hydrol., 362(3-4): 274-290.

Pagano, T.C. et al., 2014. Challenges of Operational River Forecasting. J. Hydrometeorol., 15(4): 1692-1707.

Palmer, R. et al., 2009. Weather Radar Education at the University of Oklahoma An Integrated Interdisciplinary Approach. Bulletin of the American Meteorological Society, 90(9): 1277-1282.

Pappenberger, F. et al., 2005. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrol Earth Syst Sc, 9(4): 381-393.

Pappenberger, F. et al., 2015. How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. J. Hydrol., 522: 697-713.

Pappenberger, F., Scipal, K., Buizza, R., 2008. Hydrological aspects of meteorological verification. Atmos. Sci. Lett., 9(2): 43-52.

1067    Peters-Lidard, C.D. et al., 2015. Integrated modeling of aerosol, cloud, precipitation and land

1068        processes at satellite-resolved scales. Environmental Modelling & Software, 67: 149-159.

1069    Petersen, W., Krajewski, W., 2013. Status Update on the GPM Ground Validation Iowa Flood

1070        Studies (IFloodS) Field Experiment, EGU General Assembly Conference Abstracts, pp.

1071        13345.

1072    Prat, O.P., Barros, A.P., 2010. Assessing satellite-based precipitation estimates in the Southern

1073        Appalachian mountains using rain gauges and TRMM PR. Advances in Geosciences, 25:

1074        143-153.

1075    Rabuffetti, D., Barbero, S., 2005. Operational hydro-meteorological warning and real-time flood

1076        forecasting: the Piemonte Region case study. Hydrol Earth Syst Sc, 9(4): 457-466.

1077    Rabuffetti, D., Ravazzani, G., Corbari, C., Mancini, M., 2008. Verification of operational

1078        Quantitative Discharge Forecast (QDF) for a regional warning system - the AMPHORE

1079        case studies in the upper Po River. Nat Hazard Earth Sys, 8(1): 161-173.

1080    Rakovec, O., Weerts, A.H., Hazenberg, P., Torfs, P.J.J.F., Uijlenhoet, R., 2012. State updating of

1081        a distributed hydrological model with Ensemble Kalman Filtering: effects of updating

1082        frequency and observation network density on forecast accuracy. Hydrol Earth Syst Sc,

1083        16(9): 3435-3449.

1084    Rakovec, O., Weerts, A.H., Sumihar, J., Uijlenhoet, R., 2015. Operational aspects of

1085        asynchronous filtering for flood forecasting. Hydrol. Earth Syst. Sci., 19(6): 2911-2924.

1086    Randrianasolo, A., Thirel, G., Ramos, M.H., Martin, E., 2014. Impact of streamflow data

1087        assimilation and length of the verification period on the quality of short-term ensemble

1088        hydrologic forecasts. J. Hydrol., 519: 2676-2691.

1089    Reed, S., Schaake, J., Zhang, Z., 2007. A distributed hydrologic model and threshold frequency-

1090        based method for flash flood forecasting at ungauged locations. J. Hydrol., 337(3-4): 402-

1091        420.

1092    Rossa, A. et al., 2011. The COST 731 Action: A review on uncertainty propagation in advanced

1093        hydro-meteorological forecast systems. Atmos. Res., 100(2-3): 150-167.

1094    Rotach, M.W. et al., 2012. Uncertainty propagation for flood forecasting in the Alps: different

1095        views and impacts from MAP D-PHASE. Nat Hazard Earth Sys, 12(8): 2439-2448.

1096    Ryu, D., Crow, W.T., Zhan, X.W., Jackson, T.J., 2009. Correcting Unintended Perturbation

1097        Biases in Hydrologic Data Assimilation. J. Hydrometeorol., 10(3): 734-750.

1098    Sakov, P., Evensen, G., Bertino, L., 2010. Asynchronous data assimilation with the EnKF. Tellus

1099        A, 62(1): 24-29.

1100    Salamon, P., Feyen, L., 2009. Assessing parameter, precipitation, and predictive uncertainty in a

1101        distributed hydrological model using sequential data assimilation with the particle filter. J.

1102        Hydrol., 376(3-4): 428-442.

1103    Schaake, J.C., Hamill, T.M., Buizza, R., Clark, M., 2007. The hydrological ensemble prediction

1104        experiment. Bulletin of the American Meteorological Society, 88(10): 1541-+.

1105    Seo, D.J., Koren, V., Cajina, N., 2003. Real-time variational assimilation of hydrologic and

1106        hydrometeorological data into operational hydrologic forecasting. J. Hydrometeorol., 4(3):

1107        627-641.

1108    Shi, J.J. et al., 2014. Implementation of an aerosol-cloud-microphysics-radiation coupling into

1109        the NASA unified WRF: Simulation results for the 6-7 August 2006 AMMA special

1110        observing period. Quarterly Journal of the Royal Meteorological Society, 140(684):

1111        2158-2175.

1112    Shrestha, D.L., Robertson, D.E., Wang, Q.J., Pagano, T.C., Hapuarachchi, H.A.P., 2013.

1113        Evaluation of numerical weather prediction model precipitation forecasts for short-term

1114        streamflow forecasting purpose. Hydrol Earth Syst Sc, 17(5): 1913-1931.

1115    Sorooshian, S., Dracup, J.A., 1980. Stochastic Parameter-Estimation Procedures for Hydrologic

1116        Rainfall-Runoff Models - Correlated and Heteroscedastic Error Cases. Water Resources

1117        Research, 16(2): 430-442.

1118    Tao, J., Barros, A., 2014a. Coupled prediction of flood response and debris flow initiation during

1119        warm-and cold-season events in the Southern Appalachians, USA. Hydrol Earth Syst Sc,

1120        18(1): 367-388.

1121    Tao, J., Barros, A., 2014b. The Integrated Precipitation and Hydrology Experiment. Part I:

1122        Quality High-Resolution Landscape Attributes Datasets. Report EPL-2013-IPHEX-

1123        H4SE-1, EPL/Duke University (Pub.): 60.

1124    Tao, J., Barros, A., 2014c. The Integrated Precipitation and Hydrology Experiment. Part II:

1125        Atmospheric Forcing and Topographic Corrections. Report EPL-2013-IPHEX-H4SE-2,

1126        EPL/Duke University (Pub.): 80.

1127    Tao, J., Barros, A.P., 2013. Prospects for flash flood forecasting in mountainous regions - An

1128        investigation of Tropical Storm Fay in the Southern Appalachians. J. Hydrol., 506: 69-89.

1129    Taramasso, A.C., Gabellani, S., Parodi, A., 2005. An operational flash-flood forecasting chain

1130        applied to the test cases of the EU project HYDROPTIMET. Nat Hazard Earth Sys, 5(5):

1131        703-710.

1132    Thielen, J., Bartholmes, J., Ramos, M.H., de Roo, A., 2009. The European Flood Alert System -

1133        Part 1: Concept and development. Hydrol Earth Syst Sc, 13(2): 125-140.

1134    Vasiloff, S.V. et al., 2007. Improving QPE and very short term QPF: An initiative for a

1135        community-wide integrated approach. Bulletin of the American Meteorological Society,

1136        88(12): 1899-1911.

1137    Verbunt, M., Walser, A., Gurtz, J., Montani, A., Schar, C., 2007. Probabilistic flood forecasting

1138        with a limited-area ensemble prediction system: Selected case studies. J. Hydrometeorol.,

1139        8(4): 897-909.

1140    Versini, P.A., Berenguer, M., Corral, C., Sempere-Torres, D., 2014. An operational flood

1141        warning system for poorly gauged basins: demonstration in the Guadalhorce basin

1142        (Spain). Nat Hazards, 71(3): 1355-1378.

1143    Vincendon, B., Ducrocq, V., Nuissier, O., Vie, B., 2014. Real-time  hydro-meteorological

1144        ensemble forecasting  during the HYMEX SOP1 European Meteorological Society.

1145    Vrugt, J.A., Gupta, H.V., Nuallain, B.O., 2006. Real-time data assimilation for operational

1146        ensemble streamflow forecasting. J. Hydrometeorol., 7(3): 548-565.

1147    Wanders, N., Karssenberg, D., de Roo, A., de Jong, S.M., Bierkens, M.F.P., 2014. The suitability

1148        of remotely sensed soil moisture for improving operational flood forecasting. Hydrol

1149        Earth Syst Sc, 18(6): 2343-2357.

1150    Welles, E., Sorooshian, S., Carter, G., Olsen, B., 2007. Hydrologic verification - A call for action

1151        and collaboration. Bulletin of the American Meteorological Society, 88(4): 503-+.

1152    Werner, M., Cranston, M., Harrison, T., Whitfield, D., Schellekens, J., 2009. Recent

1153        developments in operational flood forecasting in England, Wales and Scotland. Meteorol

1154        Appl, 16(1): 13-22.

1155     Wilson, A.M., Barros, A.P., 2014. An Investigation of Warm Rainfall Microphysics in the
1156          Southern Appalachians: Orographic Enhancement via Low-Level Seeder-Feeder
1157          Interactions. Journal of the Atmospheric Sciences, 71(5): 1783-1805.

1158     Wilson, A.M., Barros, A.P., 2015. Landform Controls on Low Level Moisture Convergence and
1159          the Diurnal Cycle of Warm Season Orographic Rainfall in Middle Mountains -
1160          Observations and Modeling in the Southern Appalachians. Journal of Hydrology (In
1161          Review).

1162     Wooten, R.M. et al., 2008. Geologic, geomorphic, and meteorological aspects of debris flows
1163          triggered by Hurricanes Frances and Ivan during September 2004 in the Southern
1164          Appalachian Mountains of Macon County, North Carolina (southeastern USA).
1165          Landslides, 5(1): 31-44.

1166     Xie, X.H., Zhang, D.X., 2010. Data assimilation for distributed hydrological catchment modeling
1167          via ensemble Kalman filter. Adv. Water Resour., 33(6): 678-690.

1168     Xuan, Y., Cluckie, I.D., Wang, Y., 2009. Uncertainty analysis of hydrological ensemble
1169          forecasts in a distributed model utilising short-range rainfall prediction. Hydrol Earth
1170          Syst Sc, 13(3): 293-303.

1171     Yildiz, O., Barros, A.P., 2005. Climate Variability and Hydrologic Extremes-Modeling the
1172          Water and Energy Budgets in the Monongahela River Basin. Climate and Hydrology in
1173          Mountain Areas. Wiley.

1174     Yildiz, O., Barros, A.P., 2007. Elucidating vegetation controls on the hydroclimatology of a mid-
1175          latitude basin. J. Hydrol., 333(2-4): 431-448.

1176     Yildiz, O., Barros, A.P., 2009. Evaluating spatial variability and scale effects on hydrologic
1177          processes in a midsize river basin. Scientific Research and Essays, 4(4): 217-225.

1178    Zaitchik, B.F., Santanello, J.A., Kumar, S.V., Peters-Lidard, C.D., 2013. Representation of Soil

1179        Moisture Feedbacks during Drought in NASA Unified WRF (NU-WRF). J.

1180        Hydrometeorol., 14(1): 360-367.

1181    Zappa, M. et al., 2010. Propagation of uncertainty from observing systems and NWP into

1182        hydrological models: COST-731 Working Group 2. Atmos. Sci. Lett., 11(2): 83-91.

1183    Zappa, M., Jaun, S., Germann, U., Walser, A., Fundel, F., 2011. Superposition of three sources

1184        of uncertainties in operational flood forecasting chains. Atmos. Res., 100(2-3): 246-262.

1185    Zappa, M. et al., 2008. MAP D-PHASE: real-time demonstration of hydrological ensemble

1186        prediction systems. Atmos. Sci. Lett., 9(2): 80-87.

1187    Zhang, J. et al., 2014. Initial operating capabilities of quantitative precipitation estimation in the

1188        multi-radar multi-sensor system, 28th Conf. on Hydrology, Amer. Meteor. Soc., Atlanta,

1189        GA.

1190    Zhang, S.Q., Zupanski, M., Hou, A.Y., Lin, X., Cheung, S.H., 2013. Assimilation of

1191        Precipitation-Affected Radiances in a Cloud-Resolving WRF Ensemble Data

1192        Assimilation System. Mon. Weather Rev., 141(2): 754-772.

1193    Zoccatelli, D., Borga, M., Zanon, F., Antonescu, B., Stancalie, G., 2010. Which rainfall spatial

1194        information for flash flood response modelling? A numerical investigation based on data

1195        from the Carpathian range, Romania. J. Hydrol., 394(1-2): 148-161.

1196    Zupanski, D., Zhang, S.Q., Zupanski, M., Hou, A.Y., Cheung, S.H., 2011. A Prototype WRF-

1197        Based Ensemble Data Assimilation System for Dynamically Downscaling Satellite

1198        Precipitation Observations. J. Hydrometeorol., 12(1): 118-134.

1199

## List of Tables

## List of Figures

1285

1286

1287        Table 1 – Information about the stream gauges of the 12 forecast basins.

| Forecast Basins | Site No. | Station Name | Latitude | Longitude | HUC Code | Drainage Area(km$^2$) | Basin |
|---|---|---|---|---|---|---|---|
| 1 | 03460000 | CATALOOCHEE CREEK NEAR CATALOOCHEE, NC | 35.667500 | -83.073611 | 6010106 | 127.4 | Pigeon |
| 2 | 03455500 | WEST FORK PIGEON RIVER ABOVE LAKE LOGAN NR HAZELWOOD, NC | 35.396111 | -82.937500 | 6010106 | 71.5 | |
| 3 | 03456500 | EAST FORK PIGEON RIVER NEAR CANTON, NC | 35.461667 | -82.869722 | 6010106 | 133.4 | |
| 4 | 03439000 | FRENCH BROAD RIVER AT ROSMAN, NC | 35.143333 | -82.824722 | 6010105 | 175.9 | Upper French Broad |
| 5 | 03441000 | DAVIDSON RIVER NEAR BREVARD, NC | 35.273056 | -82.705833 | 6010105 | 104.6 | |
| 6 | 02149000 | COVE CREEK NEAR LAKE LURE, NC | 35.423333 | -82.111667 | 3050105 | 204.6 | Upper Broad |
| 7 | 02150495 | SECOND BROAD RIVER NR LOGAN, NC | 35.404444 | -81.872500 | 3050105 | 223.3 | |
| 8 | 02137727 | CATAWBA R NR PLEASANT GARDENS, NC | 35.685833 | -82.060278 | 3050101 | 326.3 | Upper Catawba |
| 9 | 02138500 | LINVILLE RIVER NEAR NEBO, NC | 35.794722 | -81.89 | 3050101 | 172.8 | |
| 10 | 02140991 | JOHNS RIVER AT ARNEYS STORE, NC | 35.833611 | -81.711944 | 3050101 | 520.6 | |
| 11 | 02111000 | YADKIN RIVER AT PATTERSON, NC | 35.990833 | -81.558333 | 3040101 | 74.6 | Upper Yadkin |
| 12 | 02111180 | ELK CREEK AT ELKVILLE, NC | 36.071389 | -81.403056 | 3040101 | 131.8 | |

1288

1289

1290 Table 2 – Data assimilation schemes tested and the associated implementation parameters, i.e.
1291 assimilation frequency (AF) and time window (TW). Three efficiency indices including NSE
1292 (Nash–Sutcliffe efficiency) (Nash and Sutcliffe, 1970), the KGE (Kling-Gupta Efficiency), and
1293 the modified KGE (Gupta et al., 2009; Kling et al., 2012) of the produced hindcast simulation are
1294 shown for each basin. In addition, the error in peak value (EPV, $m^3/s$) and the error in peak time
1295 (EPT, in minutes) are also provided. The best NSE and the used DA scheme for each basin are
1296 highlighted.

| Scheme | TW | AF | Name | Basin | NSE | KGE1 | KGE2 | EPV | EPT |
|---|---|---|---|---|---|---|---|---|---|
| EnKF | | 15min | EnKF_AF15min | B01 | 0.76 | 0.79 | 0.80 | 3.95 | -135 |
| | | | | B02 | 0.45 | 0.42 | 0.55 | 33.23 | 75 |
| | | | | B03 | 0.47 | 0.65 | 0.63 | 5.74 | 195 |
| | | 30min | EnKF_AF30min | B01 | 0.69 | 0.71 | 0.75 | 4.78 | -105 |
| | | | | B02 | 0.45 | 0.44 | 0.56 | 29.70 | 90 |
| | | | | B03 | 0.41 | 0.50 | 0.60 | 9.14 | -45 |
| | | 1hour | EnKF_AF60min | B01 | 0.61 | 0.58 | 0.65 | 5.65 | -270 |
| | | | | B02 | 0.34 | 0.34 | 0.50 | 33.92 | 90 |
| | | | | B03 | 0.19 | 0.35 | 0.47 | 12.23 | -270 |
| AEnKF | 1hr | 15min | AEnKF_AF15min_TW1hr | B01 | 0.71 | 0.65 | 0.75 | 4.09 | 15 |
| | | | | B02 | 0.06 | 0.22 | 0.41 | 30.50 | 75 |
| | | | | B03 | 0.93 | 0.94 | 0.95 | -0.88 | 30 |
| | | 30min | AEnKF_AF30min_TW1hr | B01 | 0.58 | 0.62 | 0.69 | 5.17 | -15 |
| | | | | B02 | 0.33 | 0.32 | 0.52 | 32.18 | 90 |
| | | | | B03 | 0.97 | 0.93 | 0.95 | 0.20 | 0 |
| | | 1hour | AEnKF_AF60min_TW1hr | B01 | 0.55 | 0.50 | 0.65 | 6.42 | -75 |
| | | | | B02 | 0.38 | 0.39 | 0.55 | 25.56 | 90 |
| | | | | B03 | 0.88 | 0.90 | 0.92 | -5.90 | 45 |
| | 2hr | 15min | **AEnKF_AF15min_TW2hr** | B01 | 0.79 | 0.70 | 0.81 | 3.76 | -135 |
| | | | | B02 | 0.37 | 0.38 | 0.53 | 26.36 | 90 |
| | | | | **B03** | **0.99** | 0.94 | 0.96 | 1.55 | 0 |
| | | 30min | AEnKF_AF30min_TW2hr | B01 | 0.72 | 0.75 | 0.83 | -4.19 | -30 |
| | | | | B02 | 0.52 | 0.49 | 0.64 | 28.33 | -150 |
| | | | | B03 | 0.94 | 0.92 | 0.95 | -2.33 | 0 |
| | | 1hour | AEnKF_AF60min_TW2hr | B01 | 0.79 | 0.70 | 0.80 | 3.54 | -30 |
| | | | | B02 | 0.39 | 0.47 | 0.57 | 26.92 | 90 |
| | | | | B03 | 0.76 | 0.81 | 0.85 | -1.72 | -15 |
| | 3hr | 15min | AEnKF_AF15min_TW3hr | B01 | 0.68 | 0.58 | 0.71 | 4.99 | -30 |
| | | | | B02 | 0.36 | 0.38 | 0.56 | 23.80 | 90 |
| | | | | B03 | 0.98 | 0.94 | 0.96 | 1.73 | 75 |
| | | 30min | AEnKF_AF30min_TW3hr | B01 | 0.87 | 0.78 | 0.85 | 3.44 | -45 |
| | | | | B02 | 0.29 | 0.31 | 0.50 | 33.08 | 45 |
| | | | | B03 | 0.87 | 0.82 | 0.88 | 3.66 | 45 |
| | | 1hour | AEnKF_AF60min_TW3hr | B01 | 0.57 | 0.51 | 0.66 | 6.04 | -135 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | B02 | 0.10 | 0.20 | 0.37 | 38.83 | -135 |
| | | | | B03 | 0.82 | 0.85 | 0.88 | 0.57 | 0 |
| EnKS | 1hr | 15min | **EnKS_AF15min_TW1hr** | B01 | 0.89 | 0.91 | 0.93 | 2.27 | -45 |
| | | | | **B02** | **0.71** | 0.58 | 0.72 | 22.10 | -30 |
| | | | | B03 | 0.83 | 0.76 | 0.79 | 5.20 | -15 |
| | | 30min | EnKS_AF30min_TW1hr | B01 | 0.76 | 0.74 | 0.81 | 3.72 | -180 |
| | | | | B02 | 0.17 | 0.27 | 0.38 | 39.92 | -165 |
| | | | | B03 | 0.88 | 0.80 | 0.84 | 4.67 | 30 |
| | | 1hour | EnKS_AF60min_TW1hr | B01 | 0.66 | 0.72 | 0.79 | 0.80 | -90 |
| | | | | B02 | -0.01 | 0.13 | 0.32 | 41.47 | 60 |
| | | | | B03 | 0.43 | 0.50 | 0.57 | 11.94 | -270 |
| | 2hr | 15min | **EnKS_AF15min_TW2hr** | **B01** | **0.98** | 0.95 | 0.97 | 1.45 | -15 |
| | | | | B02 | 0.67 | 0.80 | 0.77 | 0.92 | 90 |
| | | | | B03 | 0.85 | 0.76 | 0.83 | 7.85 | 15 |
| | | 30min | EnKS_AF30min_TW2hr | B01 | 0.83 | 0.70 | 0.79 | 4.45 | 0 |
| | | | | B02 | 0.57 | 0.53 | 0.62 | 26.87 | 45 |
| | | | | B03 | 0.78 | 0.74 | 0.81 | 4.18 | -15 |
| | | 1hour | EnKS_AF60min_TW2hr | B01 | 0.76 | 0.65 | 0.75 | 5.00 | -90 |
| | | | | B02 | 0.49 | 0.45 | 0.58 | 29.58 | 30 |
| | | | | B03 | 0.61 | 0.66 | 0.73 | 4.08 | -165 |
| | 3hr | 15min | EnKS_AF15min_TW3hr | B01 | 0.91 | 0.84 | 0.89 | 2.81 | -60 |
| | | | | B02 | 0.61 | 0.78 | 0.78 | 2.67 | 90 |
| | | | | B03 | 0.77 | 0.87 | 0.87 | -4.63 | 135 |
| | | 30min | EnKS_AF30min_TW3hr | B01 | 0.85 | 0.75 | 0.82 | 4.13 | -150 |
| | | | | B02 | 0.43 | 0.46 | 0.59 | 23.98 | 90 |
| | | | | B03 | 0.79 | 0.84 | 0.84 | 1.09 | 45 |
| | | 1hour | EnKS_AF60min_TW3hr | B01 | 0.81 | 0.79 | 0.85 | 2.68 | -60 |
| | | | | B02 | 0.15 | 0.32 | 0.43 | 31.36 | 90 |
| | | | | B03 | 0.52 | 0.49 | 0.64 | 14.01 | 75 |

1297

1298

Table 3 - Perturbation methods and parameters applied in this study.

| Fields | Distribution | Perturbing Approach | Parameters |
|--------|--------------|---------------------|------------|
| NU-WRF QPFs | Log-Normal, $\text{LogN}(\mu,\sigma)$ | Multiplicative | $\mu=0$ <br> $\sigma=0.5$ |
| SW Radiation | Log-Normal, $\text{LogN}(\mu,\sigma)$ | Multiplicative | $\mu=0$ <br> $\sigma=0.1$ |
| Other atmospheric forcing (LW Rad., air temp., etc.) | Normal, $N(\mu,\sigma)$ | Additive | $\mu=0$ for all fields. <br> LW: $\sigma=15$ <br> Temp: $\sigma=5$ <br> Press: $\sigma=25$ <br> SepcHumi: $\sigma=0.8\times10^{-3}$ <br> Wind: $\sigma=3$ |
| Soil Moisture | Normal, $N(\mu,\sigma)$ | Additive | $\mu=0$ <br> $\sigma=0.05\times\theta_{top}$ |
| Saturated Hydraulic conductivity | Normal, $N(\mu,\sigma)$ | Additive | $\mu=0$ <br> $\sigma=10^{-6}$ |
| Power n | Normal, $N(\mu,\sigma)$ | Additive | $\mu=0$ <br> $\sigma=1.5$ |
| Discharge observation | Normal, $N(\mu,\sigma)$ | Additive | $\mu=0$ <br> $\sigma=0.1\times Q_{obs}$ |

1300

1301

1302 Table 4 – Evaluation metrics of forecast results with 6 hour to 15 hour maximum leading time
1303 using the identified best DA scheme for each basin.

| Basins / Metrics and Max. Forecasting lead | Metric | 6hr | 9hr | 12hr | 15hr | Forecast w/o DA |
|---|---|---|---|---|---|---|
| Basin01 (Best DA: EnKS_AF15min_TW2hr) | NSE | 0.28 | 0.53 | **0.75** | 0.43 | -11.26 |
| | KGE1 | 0.5 | 0.53 | **0.77** | 0.41 | -1.29 |
| | KGE2 | 0.5 | 0.56 | **0.72** | 0.52 | -0.79 |
| | EPV | 1.87 | 5.86 | 3.12 | 6.75 | -14.20 |
| | EPT | 240 | -75 | -105 | 45 | 120 |
| Basin02 (Best DA: EnKS_AF15min_TW1hr) | NSE | **0.43** | 0.25 | -0.19 | -0.10 | -0.04 |
| | KGE1 | **0.54** | 0.48 | 0.39 | 0.29 | 0.43 |
| | KGE2 | **0.61** | 0.54 | 0.28 | 0.28 | 0.49 |
| | EPV | 6.61 | 5.75 | -17.59 | 1.12 | -40.06 |
| | EPT | 120 | 120 | 120 | 120 | 120 |
| Basin03 (Best DA: AEnKF_AF15min_TW2hr) | NSE | **0.87** | 0.78 | 0.72 | 0.51 | -13.81 |
| | KGE1 | **0.9** | 0.86 | 0.85 | 0.54 | -1.78 |
| | KGE2 | **0.9** | 0.81 | 0.86 | 0.67 | -0.95 |
| | EPV | -3.19 | -8.62 | -2.73 | 8.44 | -51.39 |
| | EPT | 0 | 75 | 0 | 30 | 75 |

1304

1305
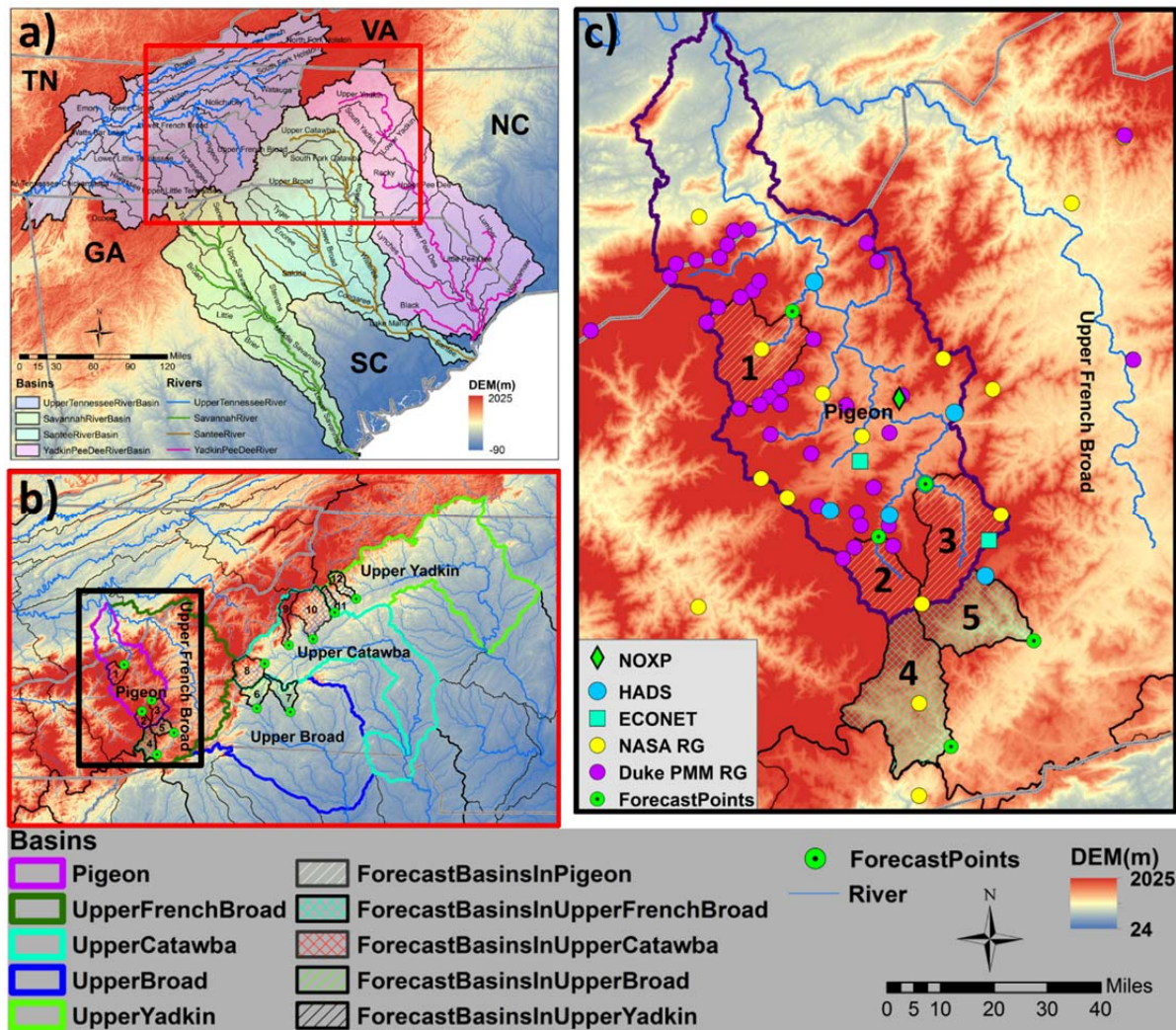
Figure 1 – The operational hydrological forecasts during the IPHEx-IOP were conducted at 12 small basins that are not limited by dam operation (labeled in panel b)), and are critical headwater catchments of the Pigeon River Basin (Basin 1-3), the Upper French Broad River Basin (Basin 4-5), the Upper Broad River Basin (Basin 6-7), the Upper Catawba River Basin (Basin 9-10) and the upper Yadkin River Basin (Basin 11-12). Green dots represent the forecasting locations which are collocated with USGS stream gauges. A dense observation network including rain gauges from NASA, Duke PMM, HADS and ECONet in the Pigeon River Basin are shown in the panel c).

1315
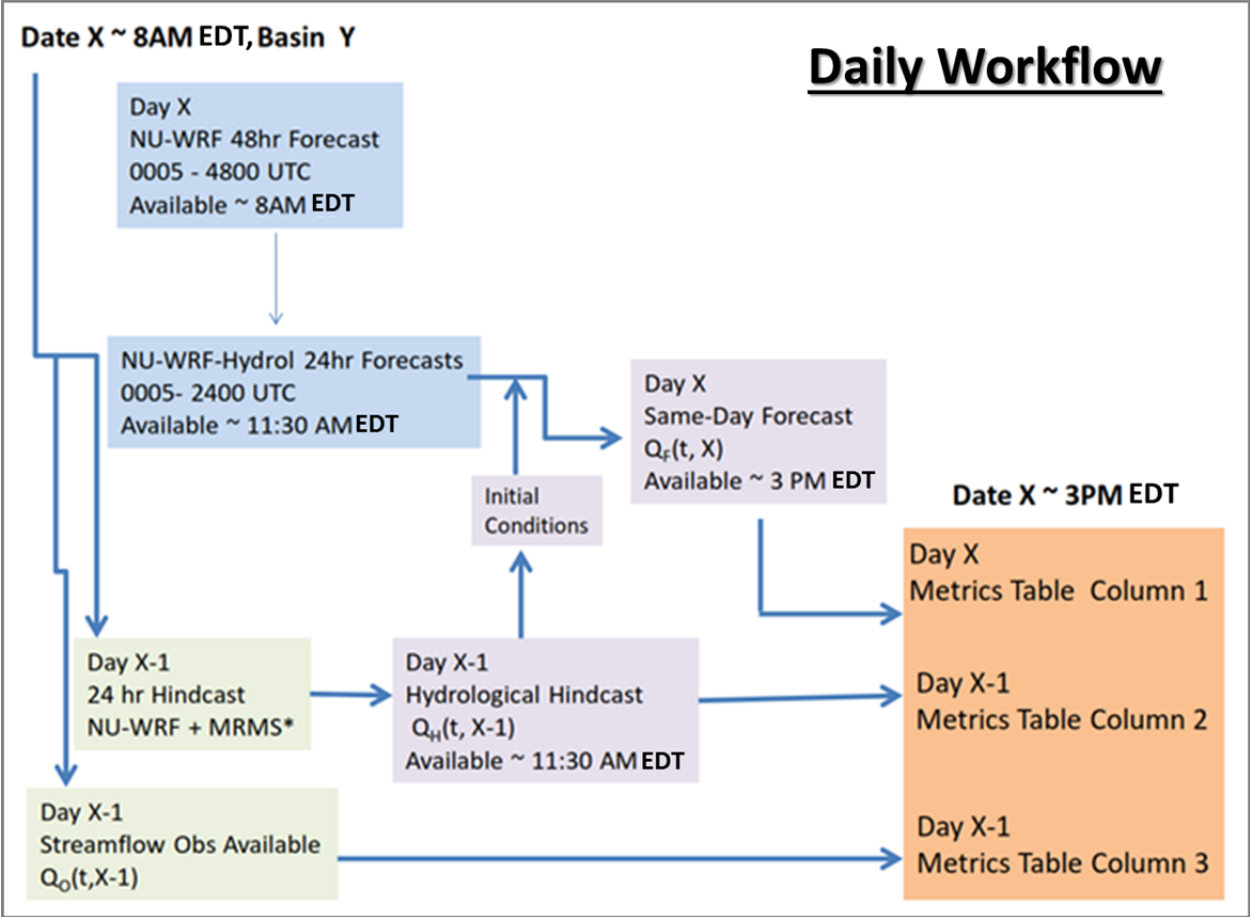


1316

1317    Figure 2 – The workflow for producing daily forecasts/hindcasts and assessment metrics at Duke

1318    (Barros et al. 2014).

1319
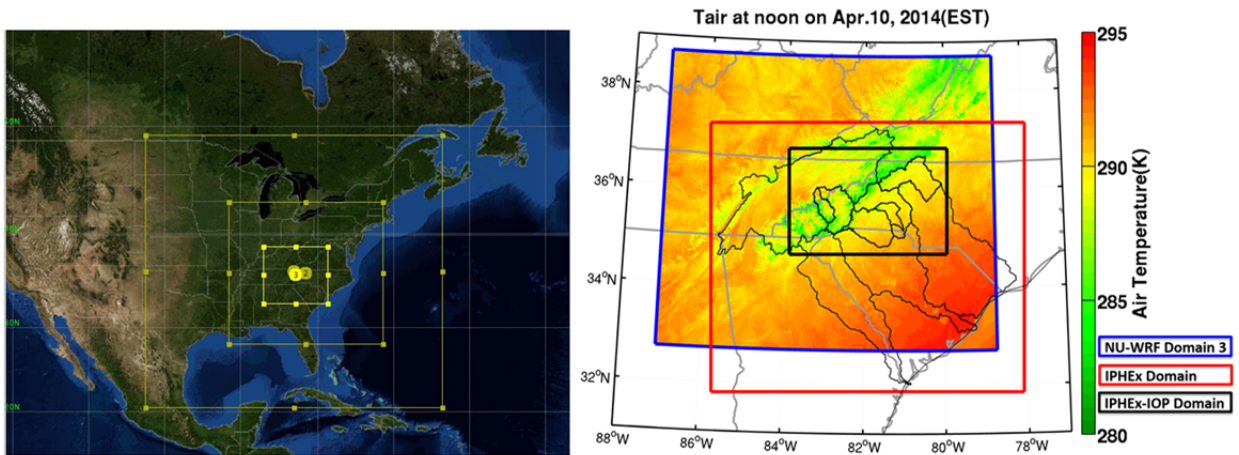
Figure 3 – The left panel shows Nu-WRF nested modeling domains during the IPHEx campaign; the right panel shows the position of the 3rd domain (the most inner) of NU-WRF, the IPHEx domain and the IPHEx-IOP domain using air temperature as an example.

1327
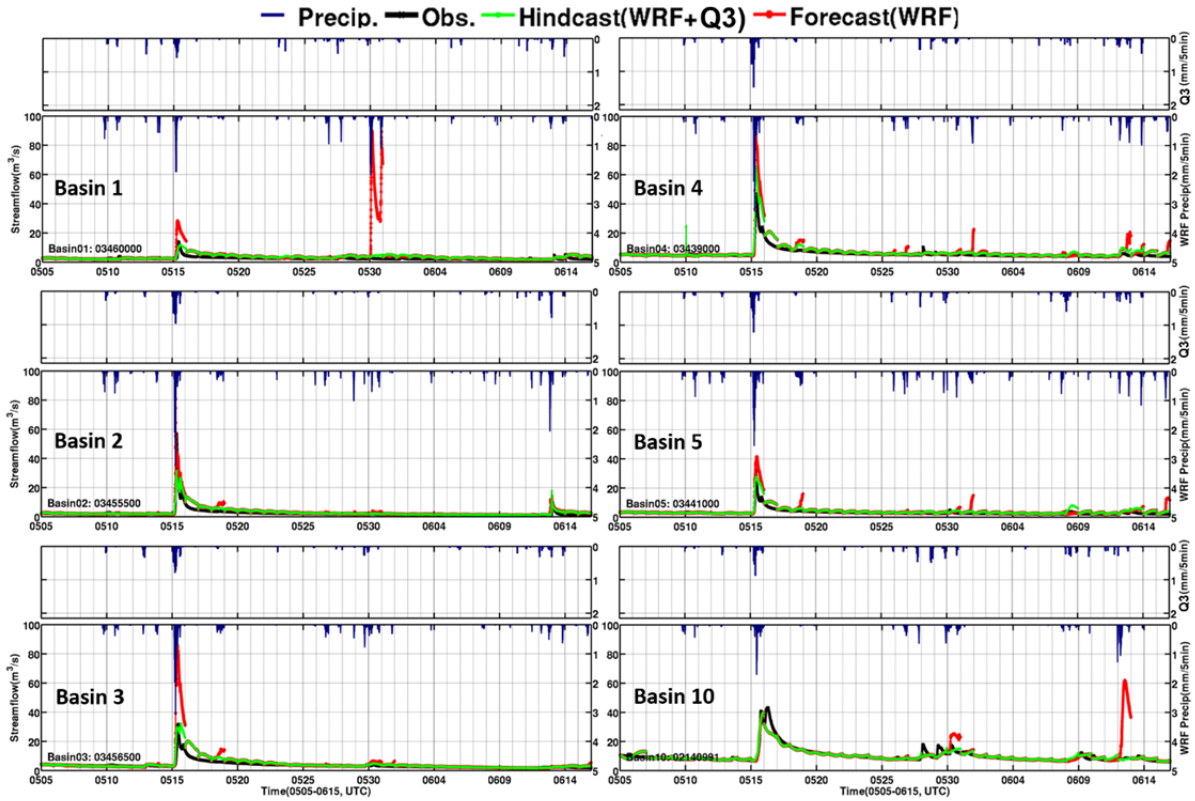
1328    Figure 4 – IPHEx-IOP Forecast/Hindcast overview (May to June 15, 2014) for Basin 1 to 5 and

1329    Basin 10, the largest basin. Dark blue represents QPE/QPF; black lines represent discharge

1330    observations; green lines are streamflow hindcast with Q3 as rainfall input and other atmospheric

1331    forcing data from Nu-WRF; red lines are streamflow forecast with all the atmospheric forcing

1332    fields from Nu-WRF.

1333

1334

Figure 5 – Daily rainfall accumulation on May 15, 2014 from ground radar-based QPEs (StageIV and Q3), satellite QPE (IMERG), QPFs from Nu-WRF operationally used in the IPHEx-IOP, and the QPFs from Nu-WRF with assimilation of conventional ground-based observations (DA CNT) and satellite-based data (DA SAT), i.e. GPM GMI and SSMIS precipitation-affected radiance. (Note the scale for QPFs from NU-WRF with DA is different from others.)
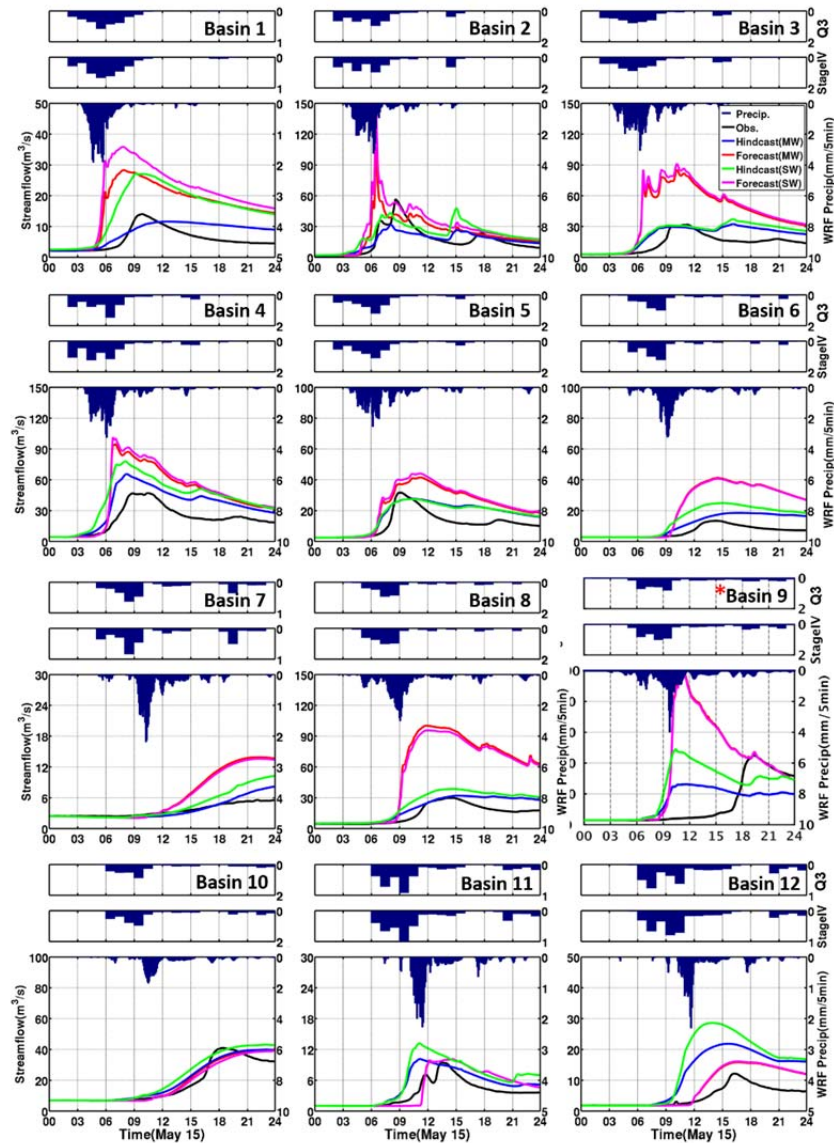
Figure 6 – IPHEx-IOP Forecast/Hindcast results for the largest event over the IPHEx (May 15, 2014) for all the basins. Dark blue represents QPE (StageIV and Q3) or QPF (Nu-WRF forecast); black lines represent discharge observations; blue and green lines are streamflow hindcasts with rainfall input from Q3 (MW) and StageIV (SW), respectively; red and pink lines are streamflow forecast with all the atmospheric forcing fields from Nu-WRF initialized using hindcast results from MW and SW, respectively.
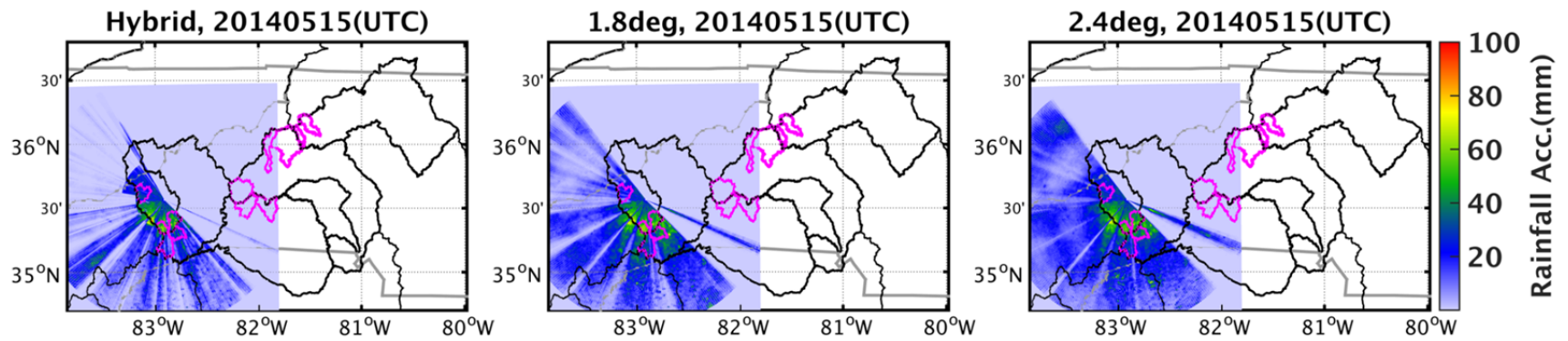
Figure 7 – Daily rainfall accumulation on May 15, 2014 from the NOAA X-band dual polarized (NOXP) radar deployed in the Pigeon River Basin. The hybrid data was produced by choosing the cleanest/lowest elevation angle for each azimuth from multiple elevation angles (from 0.5 to 8 degrees). Two other gridded NOXP data with elevation angles at 1.8 degree and 2.4 degree were also used in this study.

70
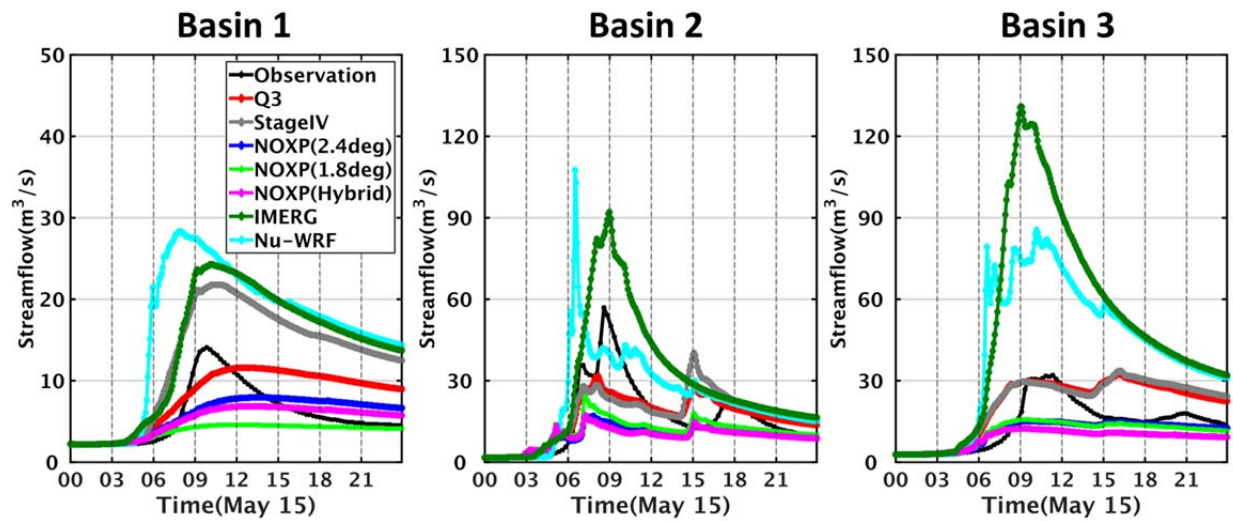
Figure 8 – Forecast/hindcast results on May 15, 2014 using multiple QPEs (Q3, StageIV, NOXP data at 1.8 degree and 2.4 degree elevation angles and the hybrid data, and IMERG) and QPF from Nu-WRF in headwater catchments in the Pigeon River Basin (Basin 1 – 3, from left to right).

1364
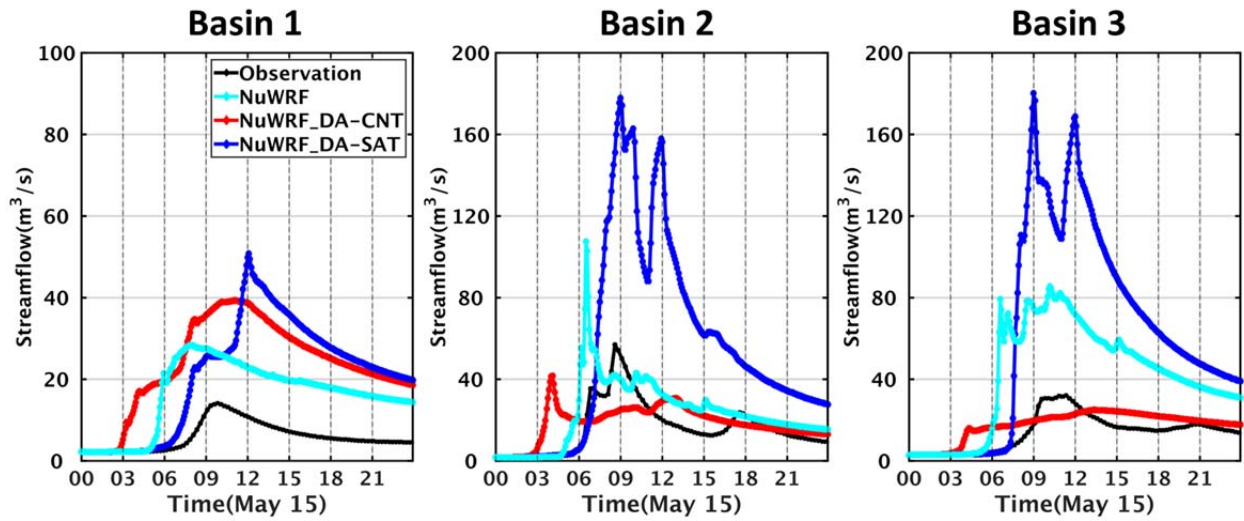
1365 Figure 9 – Forecast results on May 15, 2014 using the improved NU-WRF QPFs by assimilating

1366 conventional ground-based observations (DA-CNT), and assimilating satellite-based data (DA-

1367 SAT) (GPM GMI and SSMIS precipitation-affected radiance) also for the three headwater

1368 catchments in the Pigeon River Basin (Basin 1 – 3, from left to right).

1369

1370

1371
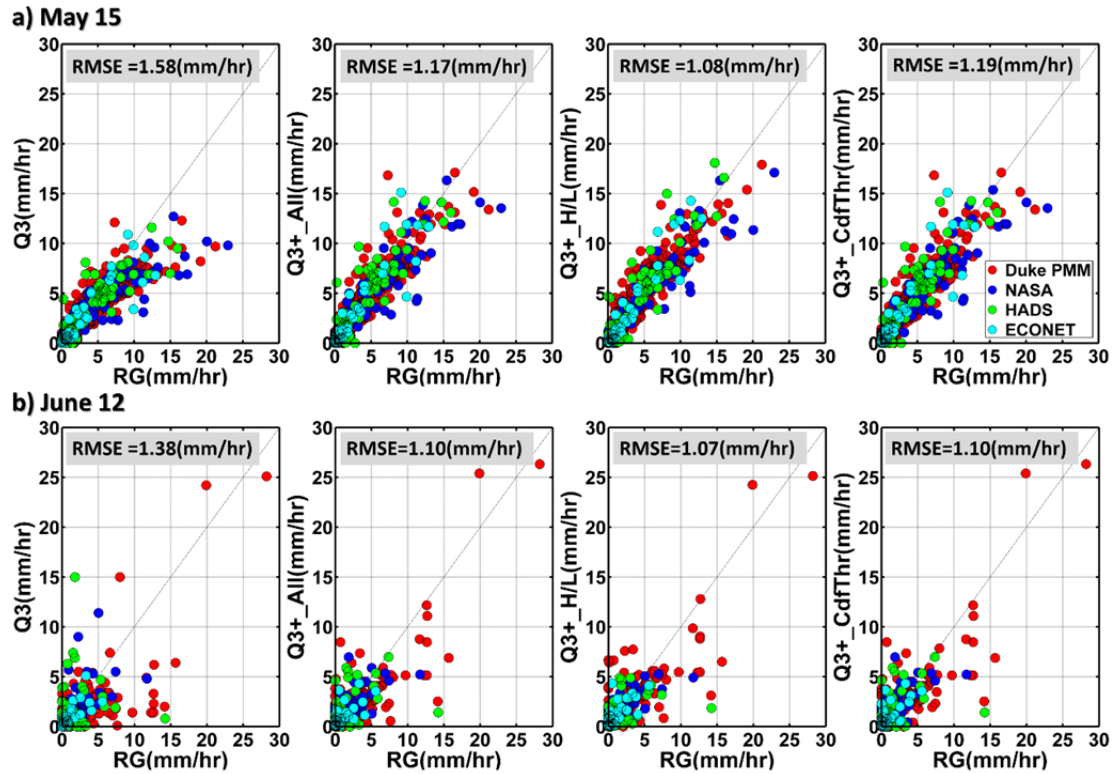
Figure 10 – Scattering comparison of the original Q3 and the adjusted Q3 data (including Q3+_All, Q3+_H/L, and Q3+_CdfThr) with observations from four raingauge networks consisting of Duke PMM gauges, NASA dual-platform, HADS and ECONet. Row a) shows the comparison for May 15 event, and row b) shows the comparison for data on June 12 (which resulted in the response on June 13).

1377
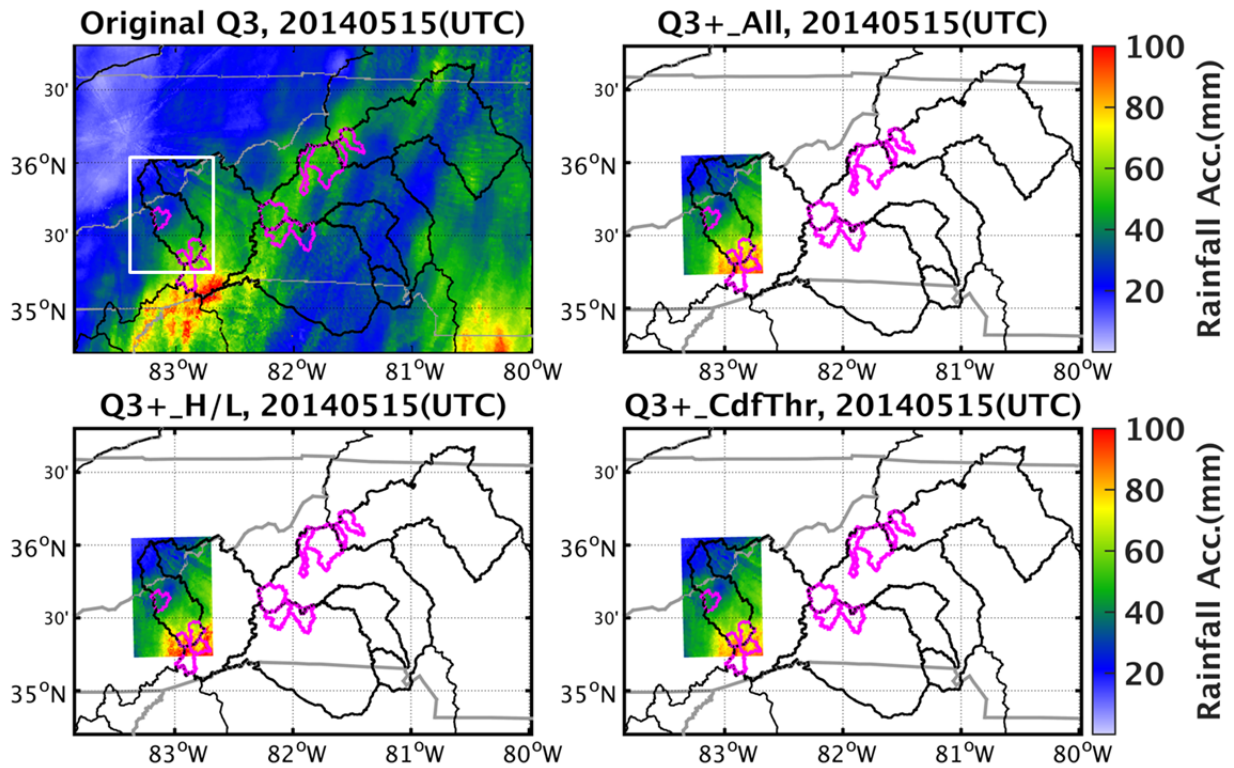
Figure 11 – Daily rainfall accumulation on May 15, 2014 from the original Q3 and the adjusted Q3 data (including Q3+_All, Q3+_H/L, and Q3+_CdfThr). Note the adjustment to Q3 data only performed in the Pigeon River Basin taking advantage of the high dense rain gauge networks.

1383
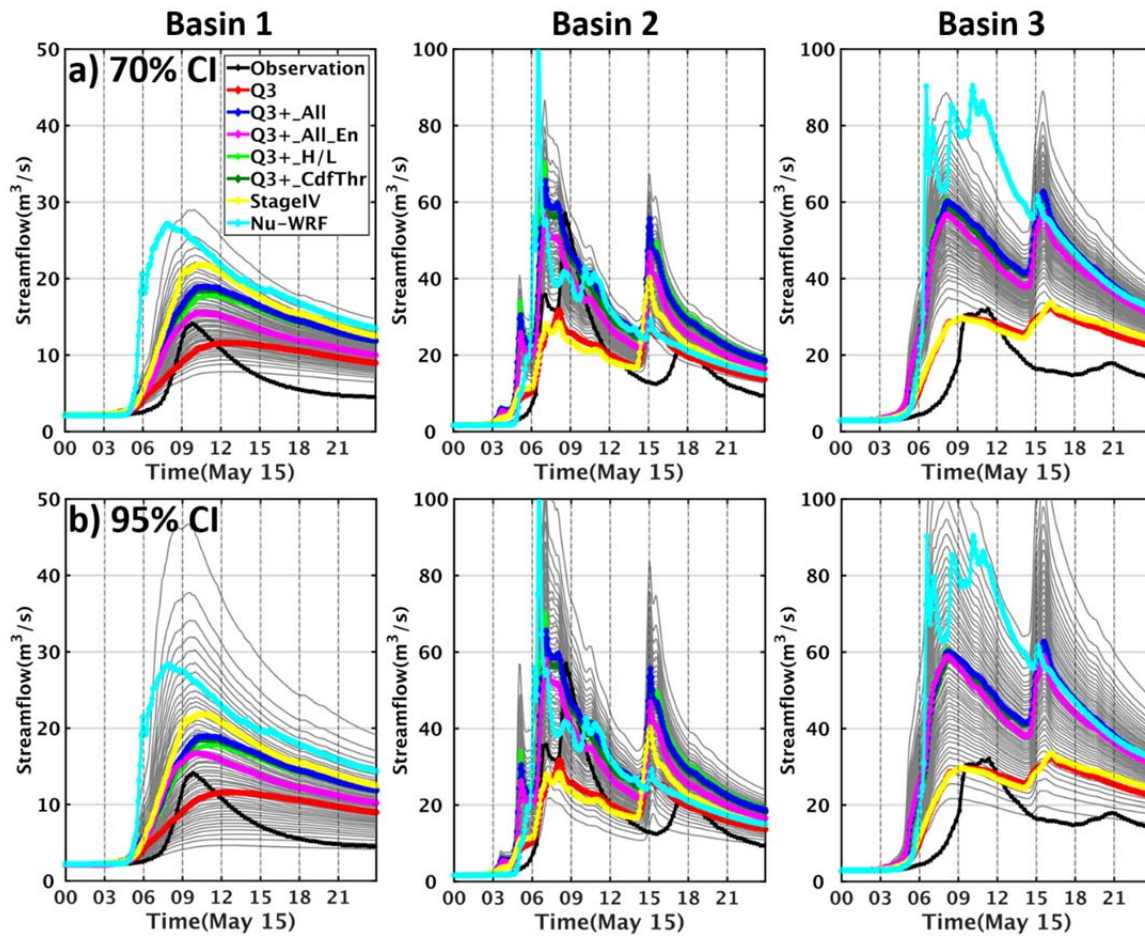
1384    Figure 12 – Forecast/hindcast results on May 15, 2014 using the original Q3 and the adjusted Q3

1385    data (Q3+_*) in headwater catchments in the Pigeon River Basin (Basin 1 – 3, from left to right).

1386    The grey lines are simulation members using 50 rainfall replicates drawn from normal

1387    distributions within 70% (row a)) and 95% (row b)) confidence interval (CI) of the regression

1388    model, explicitly representing the uncertainty associated with Q3+_All.

1389
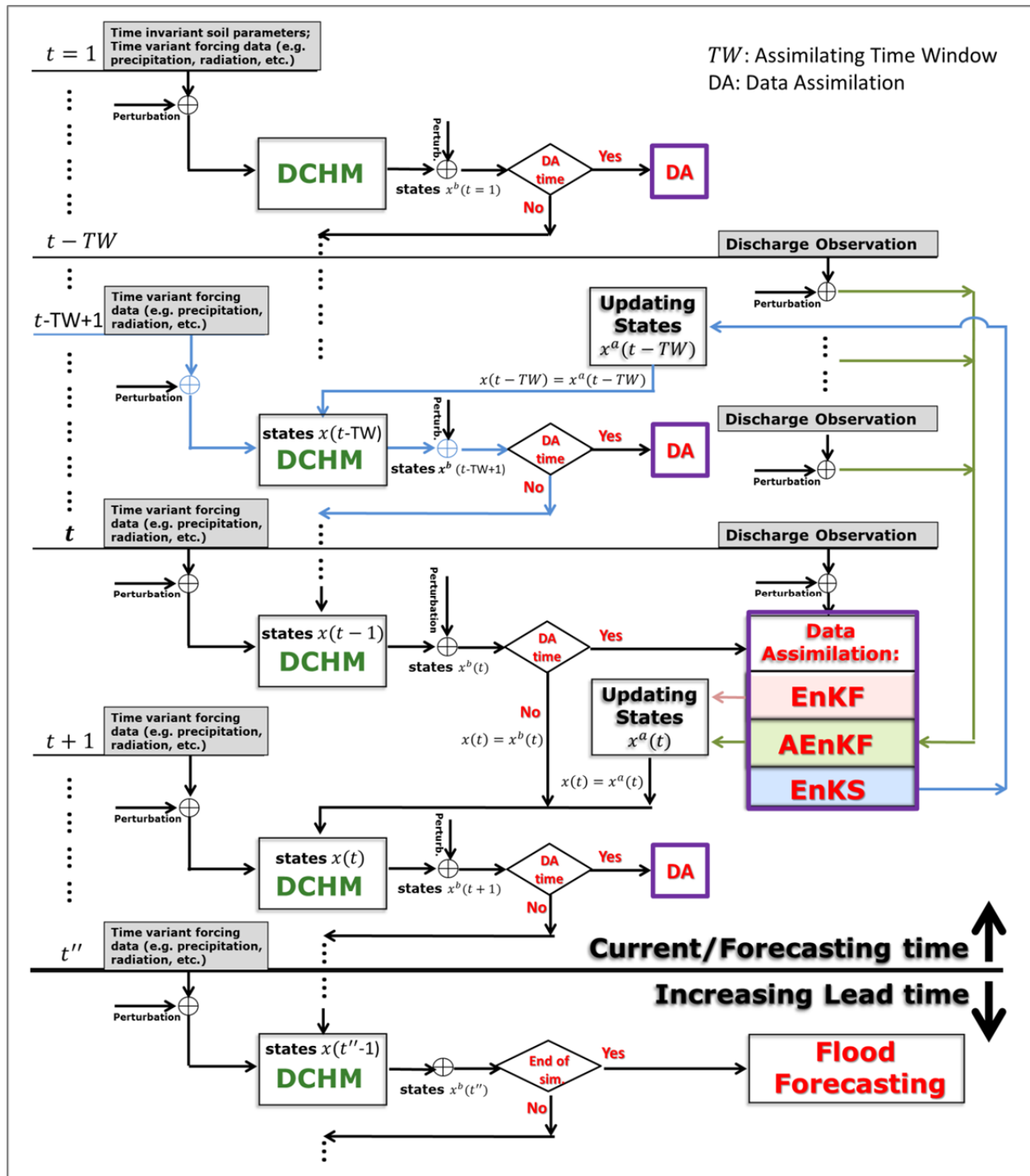
1390

1391    Figure 13 – Workflow of the hydrological Data Assimilation System (DAS) for the operational
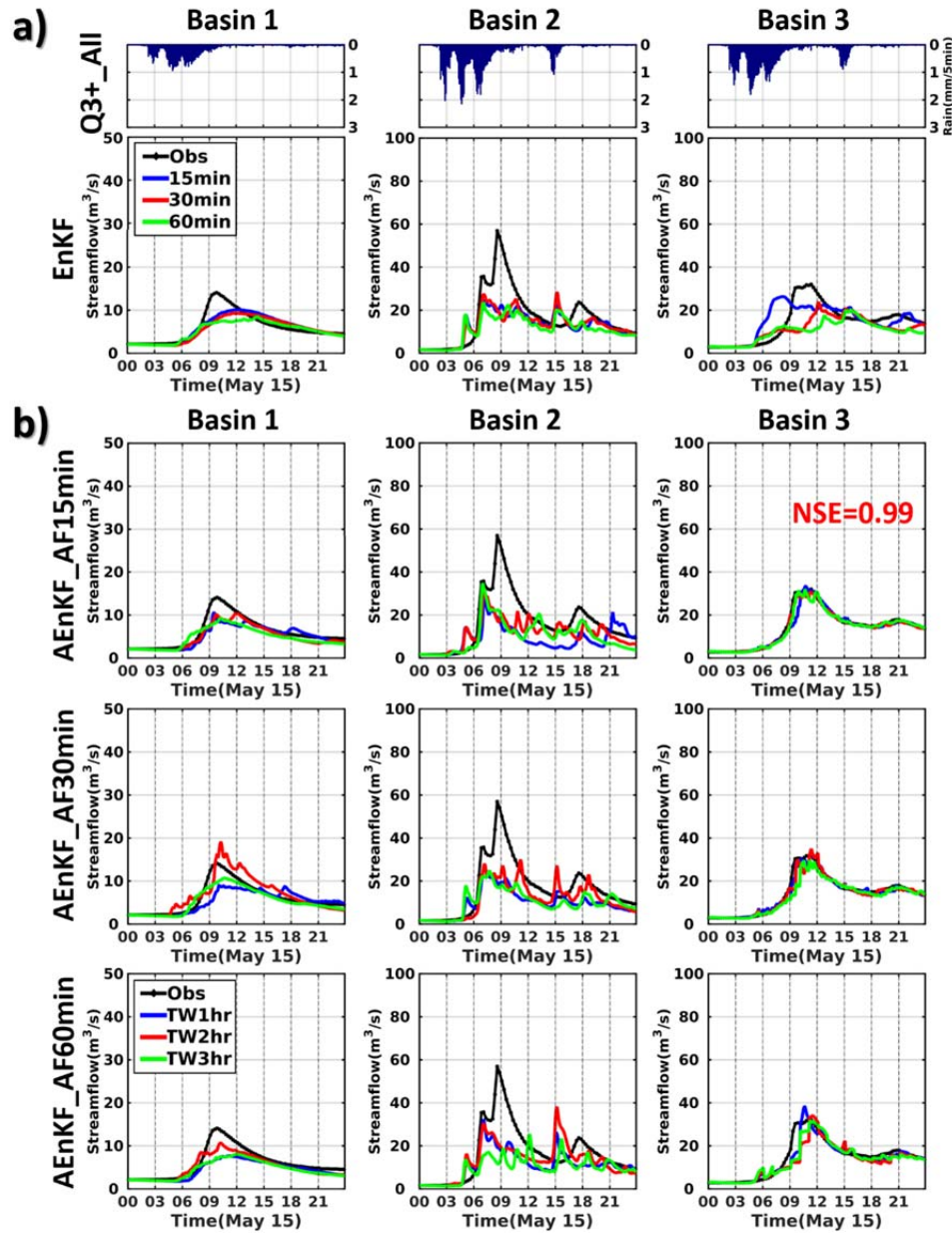
1392    flood forecast.

Figure 14 – Hindcast results assimilating discharge observations using three DA scheme, namely (a) EnKF, (b) AEnKF and (c) EnKS, with assimilation frequency (AF) from 15min, 30min to 60min, and assimilating time window (TW) from 1hr, 2hr to 3hr. Panel (d) summarizes the three schemes producing the best results indicating by NSE in Table 2. Only the ensemble members (50) of the best schemes are shown for each basin, i.e. EnKS_TW15min_TW2hr for Basin 1, EnKS_TW15min_TW1hr for Basin 2, and AEnKF_TW15min_TW2hr for Basin 3. NSEs for the best performance of DA configuration are marked in the corresponding color in the panel (see also table 2).
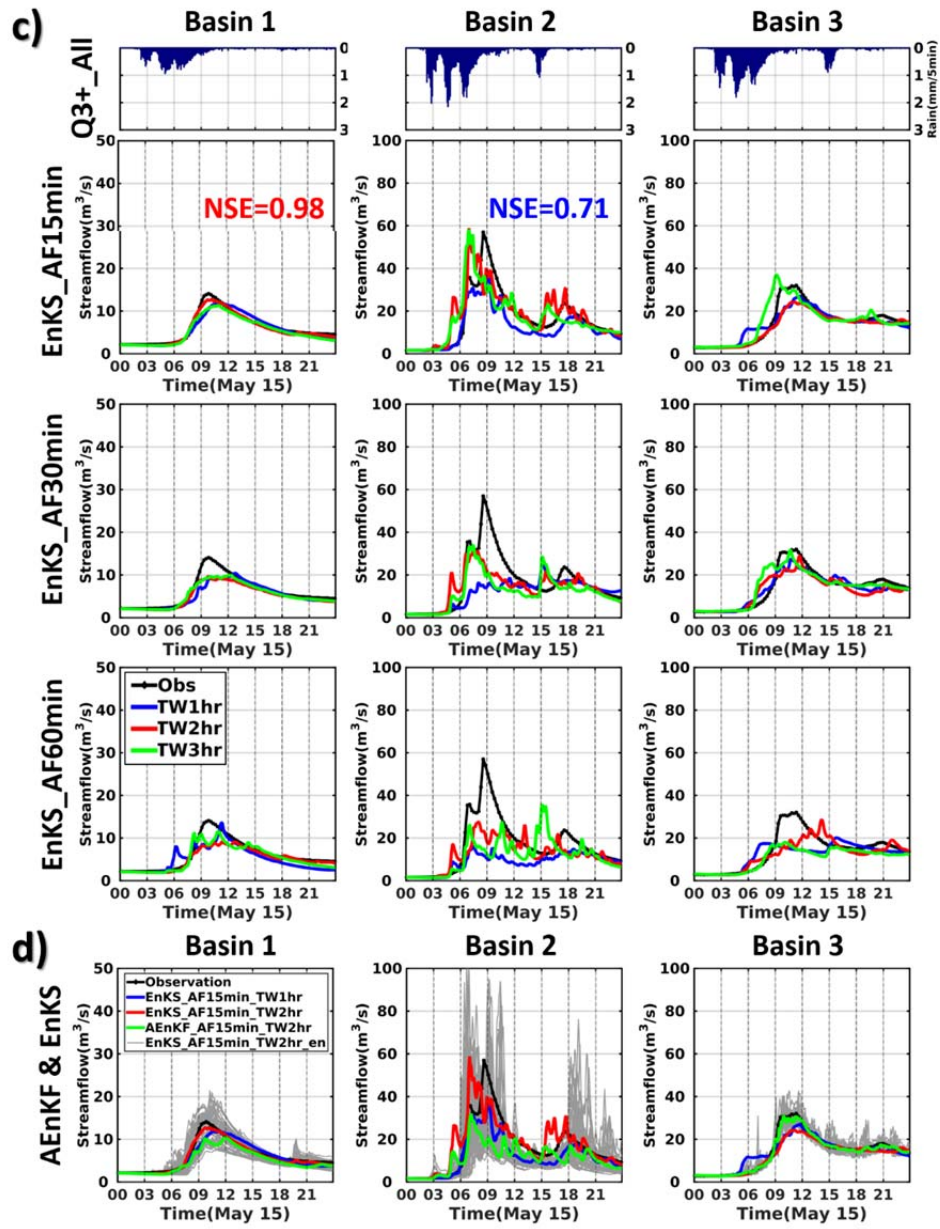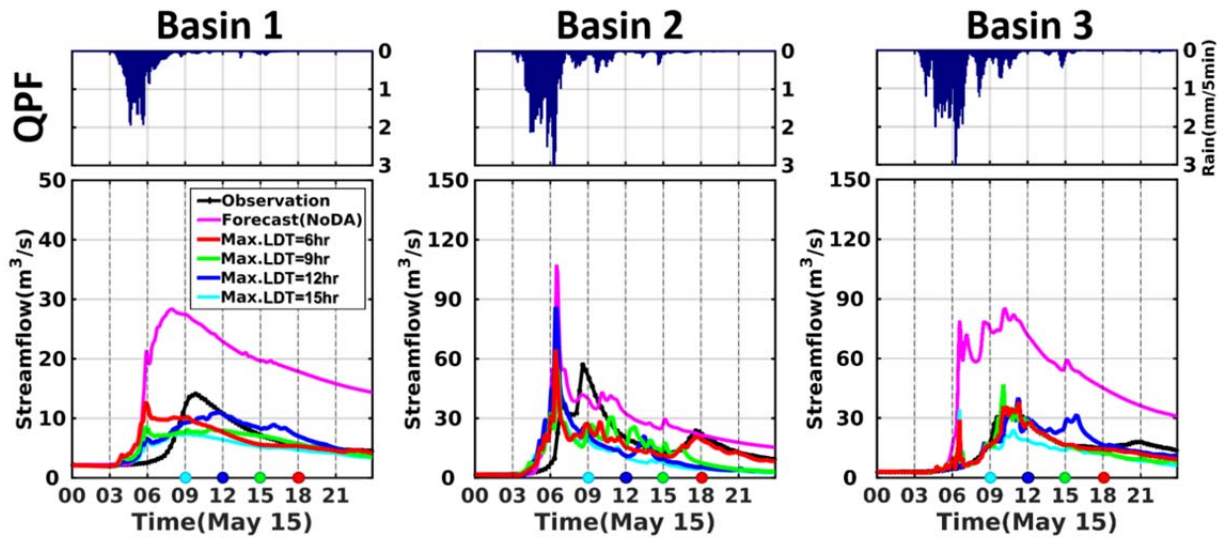
Figure 15 (continued).

1407

1408    Figure 15 – Forecast results with the best DA scheme identified for each basin (i.e.

1409    EnKS_AF15min_TW2hr for Basin 1, EnKS_AF15min_TW1hr for Basin 2, and

1410    AEnKF_AF15min_TW2hr for Basin 3) with short to longer lead times (6hr to 15hr). The time

1411    when the forecast is issued is marked on the time-axis by the dot colored corresponding to

1412    streamflow forecast. LDT means lead time.

1413

1414